

Universidad Carlos III de Madrid
Escuela Politécnica Superior



ESTUDIO DEL SECTOR INDUSTRIAL MEDIANTE MODELOS ARIMA CON DATOS ATÍPICOS

Proyecto de fin de grado
Grado en Ingeniería en Tecnologías Industriales

Juan José Villar Roldán
Leganés, Junio 2017

ESTUDIO DEL SECTOR INDUSTRIAL MEDIANTE MODELOS ARIMA CON DATOS ATÍPICOS

**Proyecto de fin de Grado
Departamento de Estadística
Grado en Ingeniería en Tecnologías Industriales**

**Alumno: Juan José Villar
Tutor: Teresa Villagarcía Casla**

Índice

1. Introducción	12
2. Estudio de la evolución del sector industrial	14
2.1 Primera etapa: 1991-1993	16
2.2 Segunda etapa: 1994-1999	18
2.3 Tercera etapa: 2000-2007	20
2.4 Cuarta etapa: 2008-2013	22
2.5 Quinta etapa: 2014-2017	24
3. Series temporales	27
3.1 Introducción	27
3.2 Características de las series temporales	27
3.3 Estructura de dependencia	28
3.3.1 Función de Autocorrelación Simple (FAS)	28
3.3.2 Función de Autocorrelación Parcial (FAP)	29
3.4 Series estacionarias	29
3.4.1 Definición	29
3.4.2 Operador de retardo	30
3.5 Procesos autorregresivos AR	30
3.5.1 Proceso autorregresivo de primer orden AR(1)	31
3.5.2 Proceso autorregresivo de orden superior AR(p)	31
3.6 Procesos de Media Móvil MA	32
3.6.1 Procesos de media móvil de primer orden MA(1)	32
3.6.2 Procesos de media móvil de orden superior MA(q)	32
3.7 Procesos ARMA(p,q)	33
3.8 Procesos Estacionales	33
3.8.1 Introducción	33
3.8.2 Series estacionales no estacionarias	35
3.8.3 Procesos ARIMA(P, D, Q)s	35
Donde P son las diferencias tomadas para el AR, Q para el MA y D las diferencias estacionales.	36
3.9 Procesos generales ARIMA (p, d, q)xP, D QS	36
4. Metodología para el análisis de series temporales	37
4.1 Identificación	37
4.2 Estimación	38
4.3 Comprobación de la Diagnósis	38

4.3.1	Análisis de los coeficientes del modelo.....	38
4.3.2	Bondad de ajuste.....	38
4.3.3	Análisis de los residuos.....	39
5.	Análisis de intervención.....	40
5.1	Variable impulso	40
5.1.1	Variable escalón: ganancia.....	43
5.2	Valores atípicos.....	44
5.2.1	Atípicos aditivos (AO)	44
5.2.2	Innovativo (IO).....	46
5.2.3	Cambio de nivel (LS)	47
6.	Análisis de las series temporales del sector industrial	48
6.1	Introducción	48
6.2	Índice de Producción Industrial (IPI).....	48
6.2.1	Definición	48
6.2.2	Datos	48
6.2.3	Análisis de datos.....	50
6.2.4	Outliers.....	53
6.3	Número total de parados	54
6.3.1	Definición	54
6.3.2	Datos	54
6.3.3	Análisis de datos.....	55
6.3.4	Outliers.....	57
6.4	Índices de Precios Industriales (IPRI).....	58
6.4.1	Definición	58
6.4.2	Datos	58
6.4.3	Análisis de Datos	59
6.4.4	Outliers.....	60
6.5	Índices de Precios de Exportación de Productos Industriales	62
6.5.1	Definición	62
6.5.2	Datos	62
6.5.3	Análisis de datos.....	63
6.5.4	Outliers.....	65
6.6	Índice de Precios de Importación de Productos Industriales	66
6.6.1	Definición	66
6.6.2	Datos	66
6.6.3	Análisis de datos.....	67

6.6.4	Outliers.....	68
6.7	Tabla de Outliers.....	69
7.	Estudio de los efectos de la crisis económica de 2008 sobre el paro con modelos ARIMA estacionales con Outliers	71
7.1	Previsiones de paro a partir de Mayo de 2017.....	72
7.2	Primer punto a analizar: Agosto 2007	74
7.3	Segundo punto a analizar: Septiembre 2008	80
7.4	Tercer punto a analizar: Junio 2011	85
8.	Conclusiones	90
8.1	Posibles trabajos futuros.....	91
9.	Bibliografía	92

Índice de figuras

Figura 1 Peso de los sectores de la Economía Española en 2016	14
Figura 2 El PIB Industrial y el PIB total de España	15
Figura 3 Evolución mensual del Índice de Producción Industrial desde enero de 1992 hasta enero de 2017.	15
Figura 4 Crecimiento mensual del Índice de Producción industrial desde Enero de 1992 hasta Enero de 2017.	16
Figura 5 Primera etapa 1991-1993	16
Figura 6 Valor añadido y ocupados de la industria sobre el total de la economía	17
Figura 7 Segunda etapa 1994-1999	18
Figura 8 Localización de la Industria Española.....	19
Figura 9 Tercera etapa 2000-2007	20
Figura 10 Gasto en I+D como porcentaje del PIB.....	21
Figura 11 Evolución de la estructura de la economía. Valor añadido 1997-2007 (Porcentaje).....	21
Figura 12 Cuarta etapa 2008-2013	22
Figura 13 Productividad por empleado en la industria española y diferencia con UE15 (2010).....	23
Figura 14 Fragmentación española en comparación con la alemana (% empleados/ tamaño de empresa)	24
Figura 15 Quinta etapa 2014-2017	24
Figura 16 Exportaciones de España. 15 principales subsectores. En Millones de euros.	25
Figura 17 Serie del Índice de Producción Industrial.....	28
Figura 18 Estructura de dependencia de una serie temporal.....	29
Figura 19 Estructura de dependencia de una serie temporal.....	31
Figura 20 AR(p).....	31
Figura 21 Ejemplo de serie con componente estacional	34
Figura 22 Modelo estacional con periodicidad anual ($s=12$)	34
Figura 23 Etapas metodología Box-Jenkins.....	37
Figura 24 Verificación. Ruido blanco.....	39
Figura 25 Atípico aditivo modelo AR (Galeano, 2011)	46
Figura 26 Atípico innovativo modelo ARIMA, (Galeano, 2011)	47
Figura 27 Cambio de nivel modelo ARIMA, (Galeano, 2011).....	47
Figura 28 Gráfica Índice de Producción Industrial	49
Figura 29 FAS y FAP del IPI previos al análisis.....	50
Figura 30 Gráfica del IPI	50
Figura 31 Gráfica IPI tras tomar una diferencia estacional	51
Figura 32 Gráfica IPI tras tomar una diferencia estacional y una regular.....	51
Figura 33 FAS y FAP tras aplicar el modelo ARIMA(1,1,2)(2,1,0) ₁₂	52
Figura 34 Gráfica Paro Total.....	54
Figura 35 FAS y FAP de Paro total	55
Figura 36 Gráfica Paro total tras quitar una diferencia regular y una estacional	55
Figura 37 Gráfica Paro total tras quitar dos diferencias regulares y una estacional	56
Figura 38 FAS y FAP tras aplicar un modelo ARIMA(0,2,1)(0,1,1) ₁₂	57
Figura 39 Gráfica Índice de Precios Industriales	58
Figura 40 FAS y FAP datos IPRI.....	59
Figura 41 FAS y FAP del IPRI tras aplicar el modelo ARIMA(1,1,0)(1,1,1) ₁₂	60
Figura 42 Gráfica del Índice de precios de exportación (IPRIX)	62

Figura 43 FAS y FAP del IPRIX.....	63
Figura 44 FAS y FAP tras aplicar un modelo ARIMA(1,1,0)	64
Figura 45 Gráfica del Índice de precios de importación de productos industriales (IPRIM).....	66
Figura 46 FAS y FAP del IPRIM	67
Figura 47 FAS y FAP del IPRIM tras aplicar el modelo ARIMA(1,1,0)	68
Figura 48 Previsiones Paro Mayo 2017.....	72
Figura 49 Gráfica Paro total (Agosto 2007).....	74
Figura 50 Previsiones paro total (Agosto 2007).....	76
Figura 51 Evolución datos reales frente a previsión (Agosto 2007)	78
Figura 52 Gráfica Paro total (Septiembre 2008)	80
Figura 53 Previsiones paro total (Septiembre 2008)	82
Figura 54 Evolución datos reales frente a previsión (Septiembre 2008)	83
Figura 55 Gráfica Paro total (Junio 2011)	85
Figura 56 Previsión paro total (Junio 2011)	87
Figura 57 Evolución datos reales frente a previsión (Junio 2011)	88

Índice de tablas

Tabla 1 Relación modelos con FAS y FAP. (Peña, 2005).....	38
Tabla 2 Outliers IPI	53
Tabla 3 Outliers Paro total	57
Tabla 4 Outliers IPRI	61
Tabla 5 Outliers IPRIX.....	65
Tabla 6 Outliers de todas las series temporales	70
Tabla 7 Porcentaje de Outliers por periodo.....	70
Tabla 8 Previsiones Paro Mayo 2017	73
Tabla 9 Outliers Paro total (Agosto 2007).....	75
Tabla 10 Predicciones Paro total (Agosto 2007)	76
Tabla 11 Datos reales frente a previsión Paro total (Agosto 2007)	77
Tabla 12 Desviación porcentual Previsto/Real Paro total (Agosto 2007)	78
Tabla 13 OutliersParo total (Septiembre 2008)	81
Tabla 14 Predicción Paro total (Septiembre 2008)	82
Tabla 15 Datos reales frente a Previsión Paro total (Septiembre 2008).....	83
Tabla 16 Desviación porcentual Paro total (Septiembre 2008)	84
Tabla 17 OutliersParo total (Junio 2011)	86
Tabla 18 Predicción Paro total (Junio 2011)	86
Tabla 19 Datos reales frente a Previsión Paro total (Junio 2011)	87
Tabla 20 Desviación porcentual Paro total (Junio 2011)	89

Series temporales

1. Introducción

El objetivo de este proyecto es estudiar el sector industrial español desde el año 1992 hasta la actualidad haciendo énfasis en las distintas crisis que ha sufrido. Se han estudiado cinco series tomadas del Instituto Nacional de Estadística. Estas cinco series son agregadas del sector industrial, pero también están disponibles las series desagregadas por sectores. Las series estudiadas son Índice de Producción Industrial (IPI), Índice de Precios Industriales (IPRI), Índice de Precios de Exportación de Precios Industriales (IPRIX) e Índice de Precios de Importación de Precios Industriales (IPRIM). También hemos estudiado el número total de parados procedente de las series de paro registrado con datos mensuales desde 1992 hasta 2017. Para realizar el análisis y obtener predicciones se ha utilizado la metodología de Box-Jenkins para modelos ARIMA (Peña, 2005). Sin embargo en ocasiones las series sufren impactos externos que les afectan y modifican sus siguientes valores. Si esto no se tiene en cuenta los modelos estimados no son correctos. El proceso de detección de estos impactos se denomina detección de valores atípicos (outliers). En este trabajo hemos explicado teóricamente varios tipos de valores atípicos. En la actualidad existen métodos para detectar su presencia en una serie temporal a través del impacto que tienen sobre la serie (Galeano, 2011). Se han modelizado las series mencionadas introduciendo la detección de atípicos y hemos encontrado una acumulación de ellos en 2008 debidos a la crisis económica.

La estructura del proyecto es la siguiente.

En el segundo capítulo se realiza un estudio de la historia reciente del sector industrial desde 1992 hasta la actualidad. Se presentan los retos a los que se enfrenta.

En el tercer capítulo se introducen los modelos ARIMA para series temporales.

En el capítulo cuarto se explica la metodología Box-Jenkins para ajustar modelos ARIMA. Esta metodología está muy contrastada tras años de uso.

En el quinto capítulo se presenta el análisis de intervención. Se han definido teóricamente varios tipos de valores atípicos y evaluado las consecuencias de no localizarlos correctamente.

En el sexto capítulo se realizará el análisis de distintos modelos de relevancia para el sector industrial español entre los que destacan el Índice de Producción Industrial, el Índice de Precios Industriales y el Paro total que han sido obtenidos del INE. La serie de datos del Paro total ha sido seleccionada por su importancia en la crisis de 2008. Además se analizarán los valores atípicos obtenidos.

En el séptimo capítulo se estudiará la crisis del 2008 a través de la serie de paro total, comparando las previsiones que se podrían haber obtenido en distintos momentos frente a los datos reales. Además se realizarán previsiones desde momento actual del paro.

Finalmente se presentan las conclusiones obtenidas de los diversos estudios realizados.

2. Estudio de la evolución del sector industrial

La economía de un país está constituida por cuatro sectores económicos básicos: agricultura, industria, construcción y servicios. La industria es el sector de la economía que transforma las materias primas en productos elaborados o semielaborados.

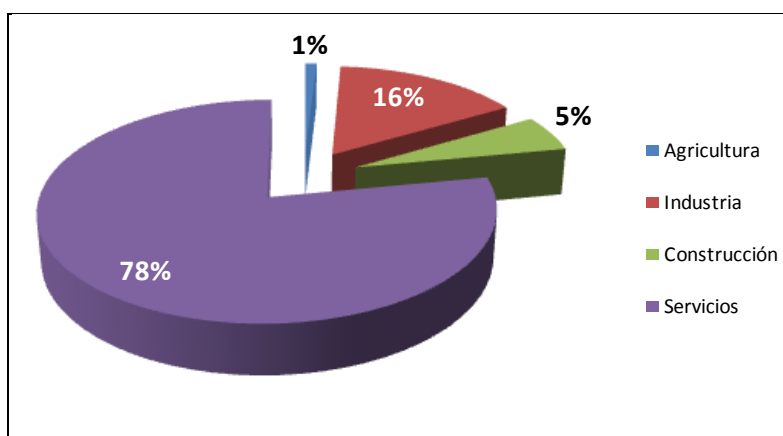


Figura 1 Peso de los sectores de la Economía Española en 2016

La Figura 1 muestra el peso los diferentes sectores en la economía española en 2016. Como puede observarse, el peso del sector industrial es del 16%, siendo el segundo en importancia tras el sector servicios.

El periodo 1990-2016 en la economía española se caracteriza por la inestabilidad y los constantes cambios. En el mismo se dan tres etapas de crisis, una a comienzos de la década de los 90, otra en el año 2000 y finalmente la sufrida en el año 2007. Dichas crisis tienen posteriormente sus recuperaciones subsiguientes. Como es bien sabido, la última crisis en 2007 ha sido la más profunda del periodo.

En únicamente 25 años la economía española, y con ella el sector industrial, se enfrenta a retos tan significativos como la entrada de España en la UE, el cambio de la peseta por el euro, el desarrollo de internet, los grandes avances en la tecnología o la aparición de la competencia procedente de países emergentes como Rusia, China o la India.

Pero hay una constante en el periodo analizado: la pérdida de peso del sector industrial en la economía española.

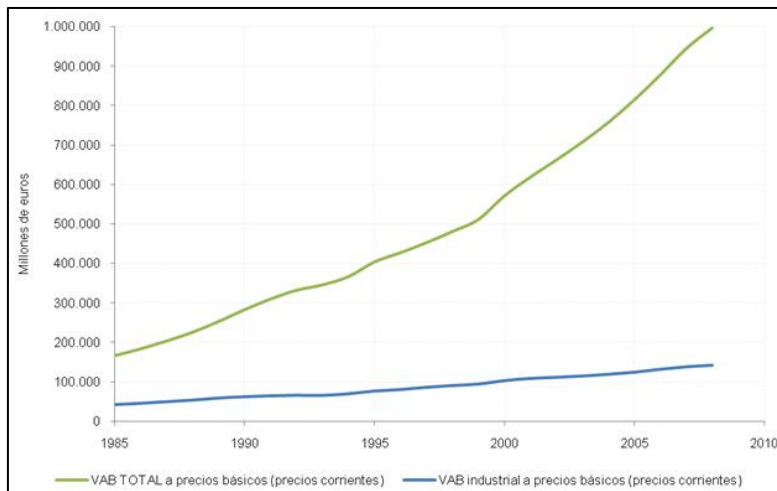


Figura 2 El PIB Industrial y el PIB total de España

La Figura 2 muestra el PIB total y el industrial de España desde 1995 a 2016. La caída de peso del sector industrial es patente.

La Figura 3 muestra la evolución mensual del Índice de Producción Industrial, una de las series que será estudiada posteriormente, desde enero de 1992 hasta enero de 2017. En ella se aprecia la magnitud de la crisis de 2007 y, en menor medida, la de 1992.

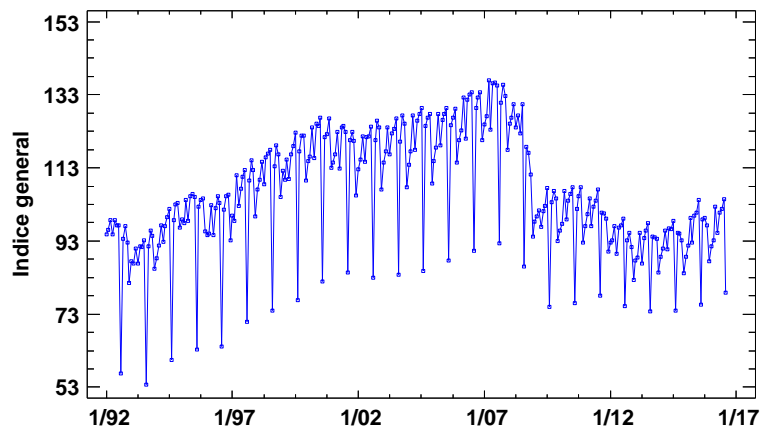


Figura 3 Evolución mensual del Índice de Producción Industrial desde enero de 1992 hasta enero de 2017.

La Figura 4, obtenida a partir de la desestacionalización de los datos de la Figura 3 muestra el crecimiento mensual del Índice de Producción industrial desde Enero de 1992 hasta Enero de 2017. En ella se aprecia claramente el tremendo impacto de la crisis económica.

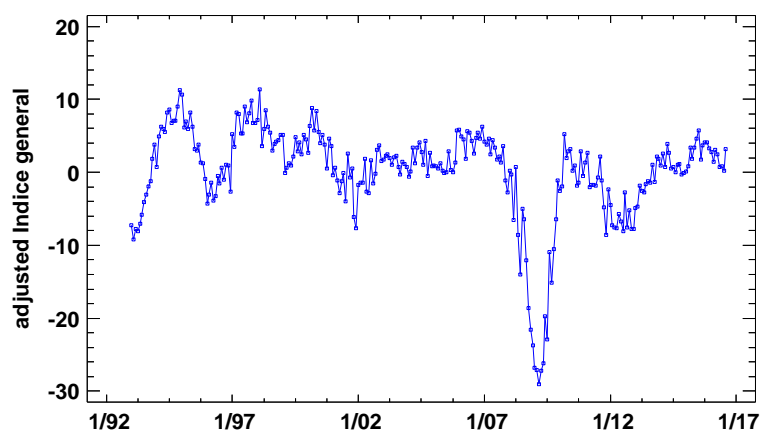


Figura 4 Crecimiento mensual del Índice de Producción industrial desde Enero de 1992 hasta Enero de 2017.

A continuación se va a analizar la evolución del sector industrial español desde la década de los 90. Para ello se consideran cinco etapas bien diferenciadas:

2.1 Primera etapa: 1991-1993

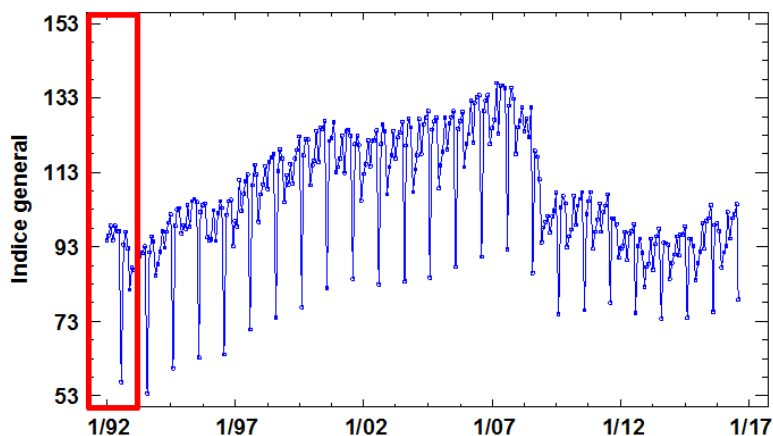


Figura 5 Primera etapa 1991-1993

La entrada de España en la Comunidad Europea en 1986 marca un cambio fundamental en la política industrial española.

La subida de los precios del petróleo provocadas por la Guerra del Golfo y los problemas inmobiliarios de Japón fueron el origen de la crisis de principios de los 90. Sin embargo, los

efectos en la economía española se retrasaron debido al esfuerzo de inversión institucional que requirieron la Expo de Sevilla, la construcción de la nueva línea del AVE, y los JJOO de 1992.

Tras la finalización de ambos acontecimientos, en 1993 la deuda pública española creció un 30%, el PIB español pasó a tasas negativas (-1%), el paro creció y la inflación alcanzó el 4,9%.

El sector industrial se vio fuertemente afectado por esta crisis. La capacidad productiva utilizada cae notablemente con respecto a la capacidad productiva instalada. Fue el peor año desde finales de la Segunda Guerra Mundial. El valor añadido aportado por la industria y el porcentaje de ocupados en el sector caen significativamente como se puede observar en la Figura 6.

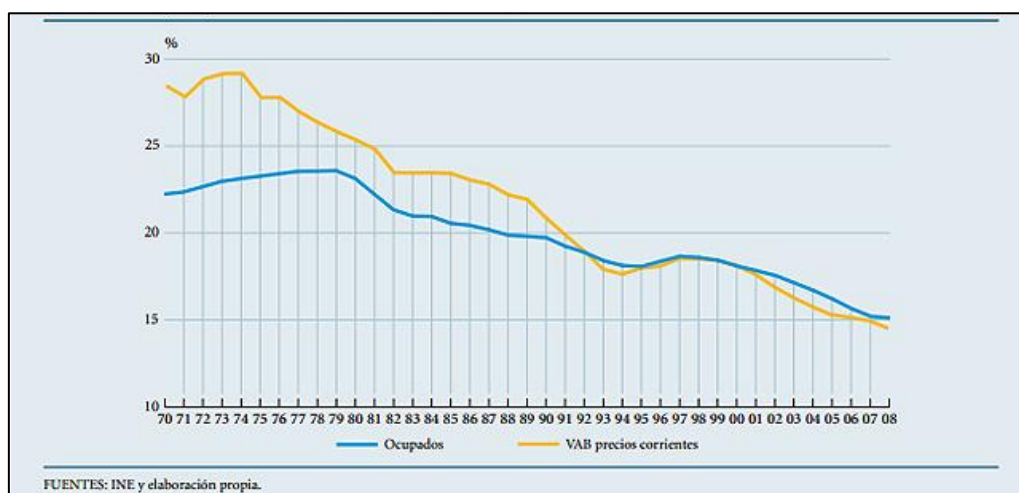


Figura 6 Valor añadido y ocupados de la industria sobre el total de la economía

En la década de 1980 se abordó una importante reconversión industrial que, tras la crisis de 1990, llevó a una recuperación económica a partir de 1994. El objetivo era poner solución a los problemas estructurales en un mundo cada vez más global y en constante cambio tecnológico. Como consecuencia, en España, se intensificaron las reconversiones para adaptarse a los mandatos comunitarios.

Esta época se caracteriza por la desaparición de aranceles, la creación de ayudas para sectores en crisis, las ayudas a las PYMES y las privatizaciones de empresas estatales.

La adaptación del tejido industrial español no fue sencilla ya que se vio obligado a competir en igualdad en el exterior –partiendo de una política proteccionista-, con falta de tamaño empresarial y desventaja tecnológica.

La inversión extranjera es uno de los factores determinantes en la modernización y competitividad de la industria española durante estos años.

2.2 Segunda etapa: 1994-1999

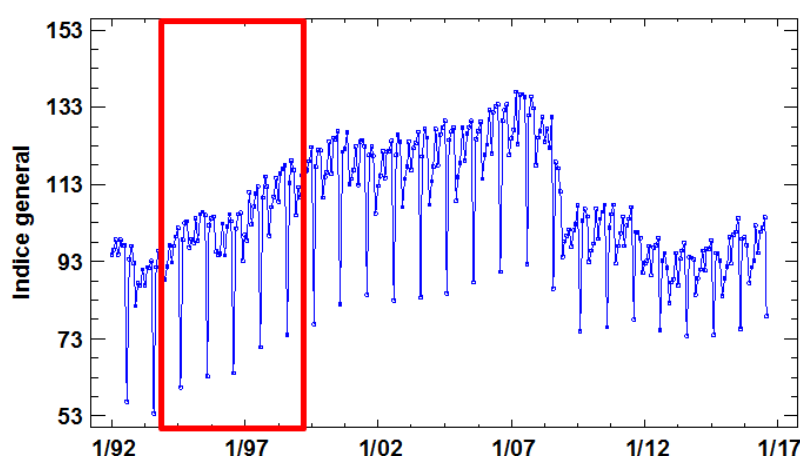


Figura 7 Segunda etapa 1994-1999

Como puede observarse en la Figura 7 la etapa 1994 a 1999 fue de crecimiento continuada. A partir de 1994 la industria es el sector más dinamizador de la economía española, contribuyendo significativamente al crecimiento del PIB y del empleo. La utilización de la capacidad productiva se eleva de nuevo a porcentajes superiores al 80%.

Las condiciones macroeconómicas, con tipos de cambio estables, bajos tipos de interés y una cierta contención salarial tuvieron una incidencia sensible tanto en la producción de bienes intermedios como en el consumo y la inversión. Destacan en este periodo las inversiones en el extranjero de grandes empresas energéticas españolas, pero también de un alto número de medianas y pequeñas empresas en zonas tradicionales de inversión desde España como son Hispanoamérica, Portugal o Marruecos. China también comienza a ser destino importante de inversiones.

El aumento de la demanda consecuencia de la globalización conllevó un fuerte crecimiento en sectores como la electrónica o la informática, junto a otros como el químico o el automóvil. Sin

embargo, otros como el textil, el mueble o la piel y el calzado sufren caídas en la producción y fuertes pérdidas de empleo. El sector alimentario se mantuvo estable.

A destacar que en 1996 se acentúa la política de privatizaciones de empresas públicas, muchas veces a través de oferta pública de acciones. Ejemplos significativos fueron Telefónica, Repsol o Endesa.

La inversión extranjera continúa su crecimiento, procedente de países de la Unión Europea en un 86% del total, principalmente de Francia, Alemania y Reino Unido.

La producción industrial se concentra en Cataluña, Comunidad Valenciana, Madrid y País Vasco y Asturias. Hay que destacar el fuerte impulso que recibió Madrid en sectores de elevado contenido en tecnología, como telecomunicaciones, equipos médicos o aeronaves. Estos sectores contribuyen ya en un 14% al total de las exportaciones industriales en 1998. La Figura 8 que muestra el porcentaje de población ocupada en la industria por provincias confirma esta distribución geográfica.

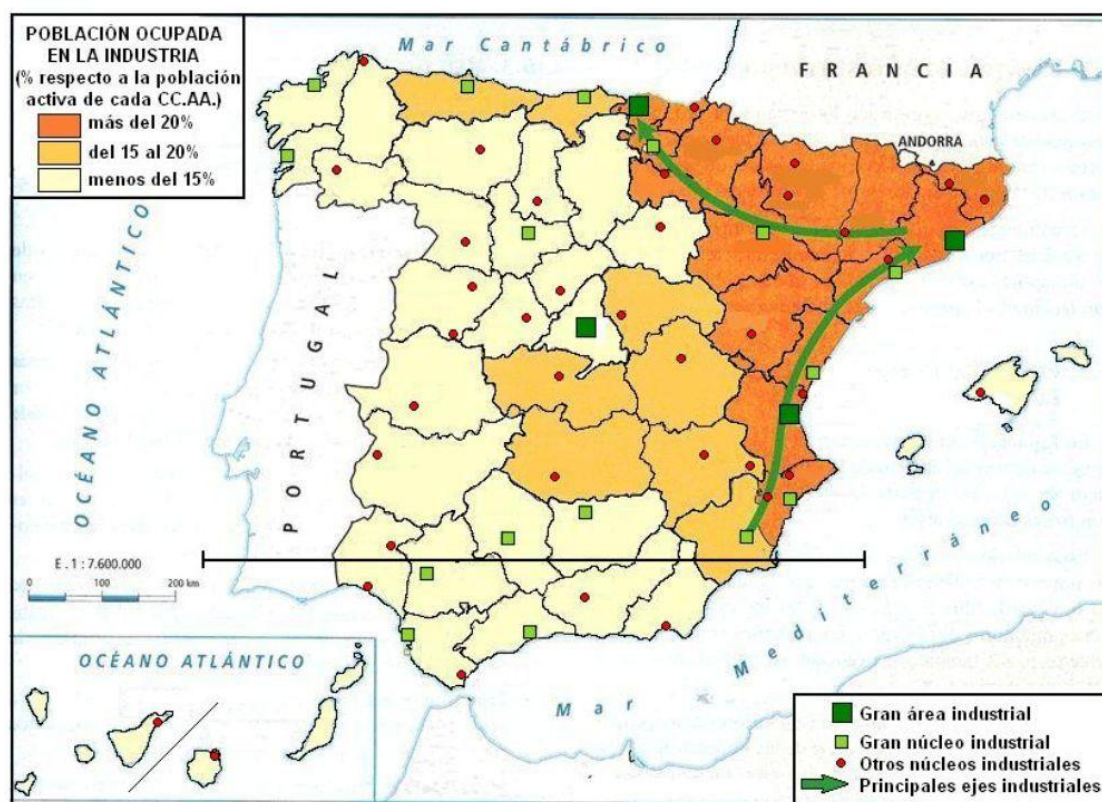


Figura 8 Localización de la Industria Española

2.3 Tercera etapa: 2000-2007

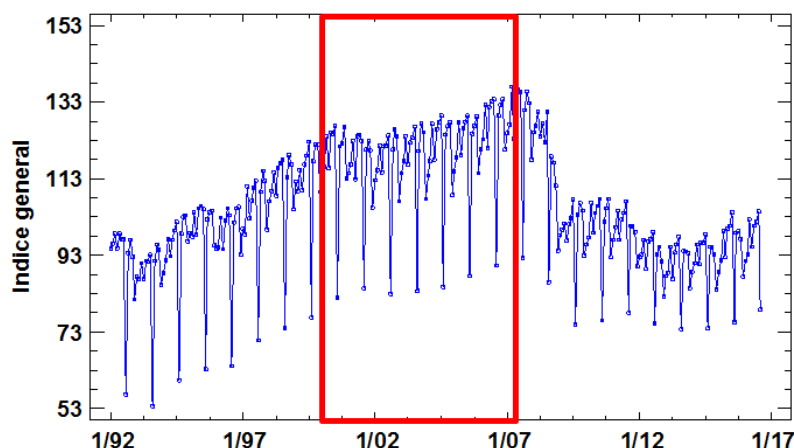


Figura 9 Tercera etapa 2000-2007

En los primeros años del SXXI se produce un pequeño ajuste en el sector industrial, seguido de un periodo de fuerte crecimiento que termina abruptamente en el 2007 con el inicio de la crisis económica. Durante ese periodo se producen hechos importantes. La ampliación de la UE a los países del Este obligó a la industria española a competir con países cuya estructura industrial era similar, pero con costes laborales mucho más bajos y con una localización geográfica más favorable a las grandes empresas del eje Suecia, Alemania y norte de Italia. Se produce una pérdida de cuota de mercado en sectores con intensidad tecnológica media y baja, e incluso en tecnología avanzada. Por ejemplo, la industria del automóvil, que representaba el 6,5% del empleo industrial y el 20% de las exportaciones fue muy sensible a esta situación. El bajo nivel de I+D en España no ayudó en esta situación, al no aportar a las empresas españolas de un nivel de diferenciación tecnológica significativa.

También a partir de principios del siglo XXI Internet comienza a representar a la vez una amenaza y una oportunidad para la empresa. La baja implantación de las TIC en la industria española, una de las más bajas de la OCDE, es otra importante desventaja competitiva.

Sin embargo, la entrada de España en el euro en el 2000 al inicio de esta etapa, con muy bajos tipos de interés, produce dos efectos que transformarán la economía española: el aumento de la demanda de viviendas y del peso del turismo. Como consecuencia, el sector industrial, a pesar de su crecimiento, continúa su pérdida de peso en la economía española a costa de los servicios y la construcción, acentuando una economía desequilibrada y poco eficiente que

dificulta su sostenibilidad. La Figura 10 muestra la evolución del gasto en I+D en España comparado con EE.UU y la UE-8. Como puede verse, el porcentaje que España dedica a I+D, ha crecido en términos absolutos, pero sigue siendo significativamente menor que el de los países desarrollados de nuestra área. La Figura 11 muestra la evolución de los sectores económicos hasta la gran crisis de 2007. Se observa el crecimiento de la construcción y la bajada de industria y sector primario, sin embargo ha de tenerse en cuenta que el sector industrial a pesar de perder peso a favor de otros sectores aumenta notablemente.

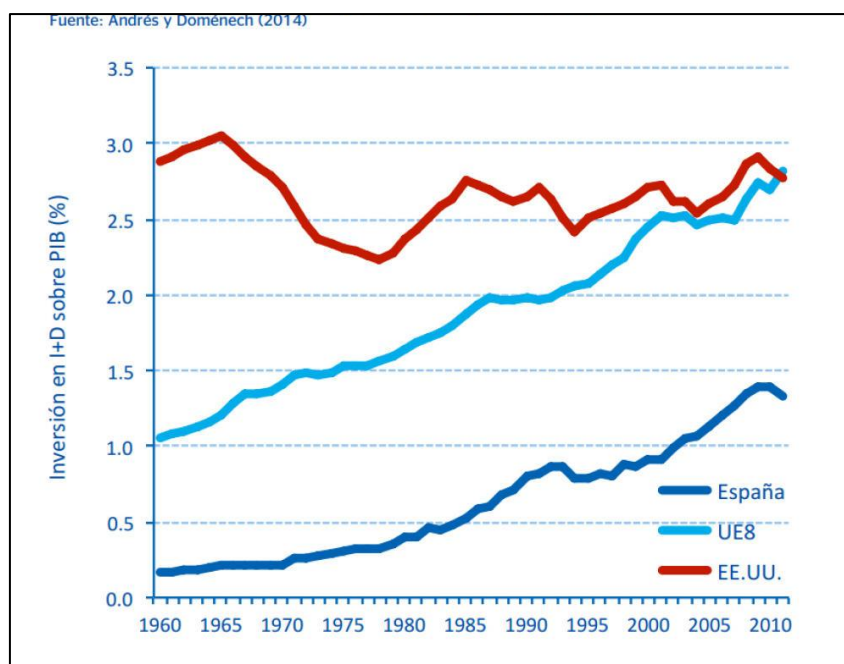


Figura 10 Gasto en I+D como porcentaje del PIB

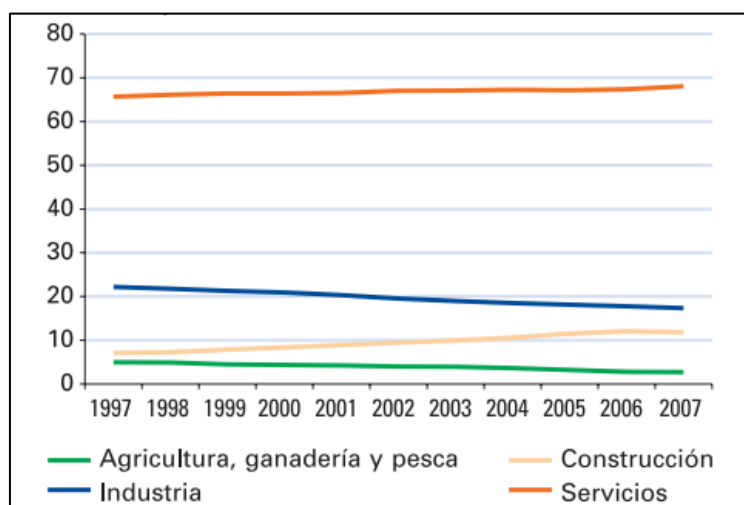


Figura 11 Evolución de la estructura de la economía. Valor añadido 1997-2007 (Porcentaje)

2.4 Cuarta etapa: 2008-2013

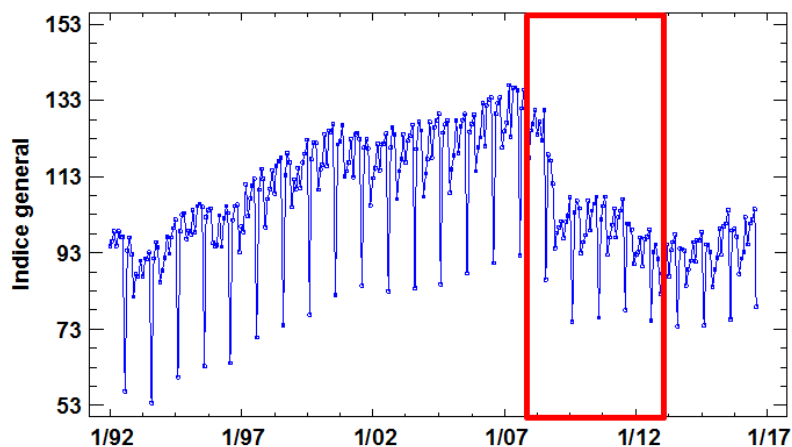


Figura 12 Cuarta etapa 2008-2013

A partir de 2008, el estallido de la burbuja inmobiliaria produce la peor crisis desde la II Guerra Mundial. El sector industrial español tuvo una enorme contracción. De hecho, en 2009 la contribución del sector industrial al PIB cayó al 14,4%, desde el 18,8% del año 2000. La cifra de negocio industrial cae en tasas del 30% comparativamente con el año anterior. Estos datos se pueden observar en la Figura 3 y en la Figura 4 que muestran la importante caída de la producción industrial española

A finales de 2011 tanto la demanda interna como las exportaciones seguían mostrando una gran debilidad. Caen a niveles de 1980 las matriculaciones de automóviles, el consumo de cemento, la construcción de nuevas viviendas y el crédito. Y la cifra de paro alcanza niveles superiores al 26%.

En 2013, el retroceso continúa, sobre todo en bienes de consumo duradero (-9,8%), energía (-7,9%) y bienes intermedios (-4,2%).

Los problemas de la industria española en esta etapa se pueden resumir en:

A) Baja Productividad

Aunque los costes laborales por hora trabajada eran significativamente bajos, el problema era una productividad menor que en los principales países competidores, por lo que la ventaja de los costes laborales quedaba diluida. Un factor importante, pero no el único, que afecta a la menor productividad eran tasas de absentismo del 8%, que más que duplicaban las de la UE.

Así, la productividad está un 34,1% por debajo de la media de la UE en el entorno del año 2010. Destaca la baja productividad en la industria farmacéutica, química o equipo eléctrico:

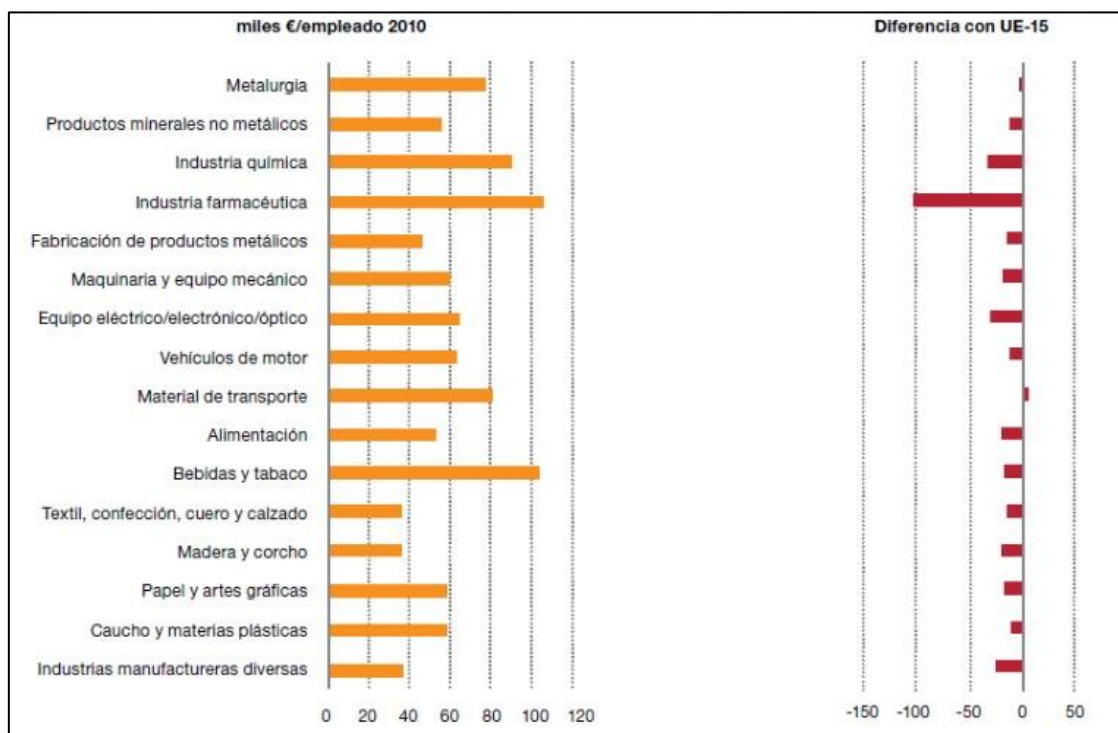


Figura 13 Productividad por empleado en la industria española y diferencia con UE15 (2010)

B) Costes de la energía

Otro hecho que dificultaba la competitividad eran los costes de la energía. España es un país que importa insumos energéticos y las empresas españolas pagan la quinta electricidad más cara de la UE por la falta de interconexiones con Europa, los impuestos y subvenciones en la tarifa o el desarrollo de renovables poco maduras. Este problema ha llevado a las empresas a buscar energías alternativas, mejorar los procesos y aumentar la eficiencia renovando maquinaria y equipos.

C) Baja inversión en I+D

Ya comentado en la etapa anterior. Tan sólo en el sector del transporte la inversión en I+D es superior al promedio europeo. Esta baja inversión implica un sector industrial con menor capacidad tecnológica y, por tanto, menor dinamismo en la exportación y mayor dependencia de las importaciones y la energía.

D) Pequeño tamaño del sector industrial

El tejido industrial español continuaba (y continúa) formado por pequeñas empresas, que aportan más del 70% del empleo (en Alemania el 50%). El problema es que aunque el número de empresas industriales en ambos países era similar, la aportación de valor añadido bruto era cuatro veces inferior en el caso español. La fragmentación dificulta la competitividad, debido a la insuficiente capacidad productiva para enfrentar proyectos de alto volumen de recursos, impide aprovechar economías de escala, dificulta el acceso a líneas de crédito y limita los recursos para I+D.

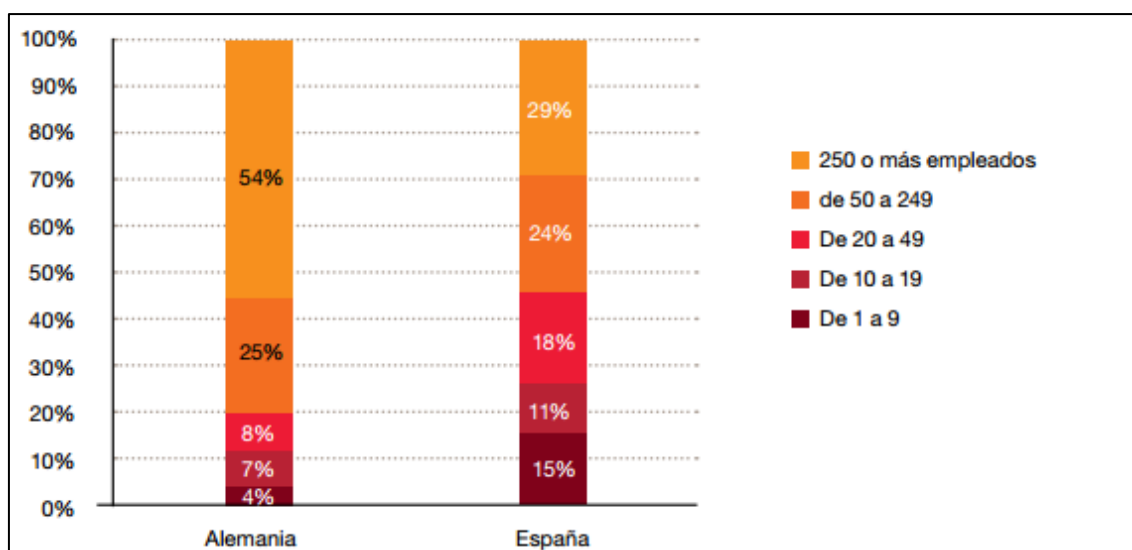


Figura 14 Fragmentación española en comparación con la alemana (% empleados/ tamaño de empresa)

2.5 Quinta etapa: 2014-2017

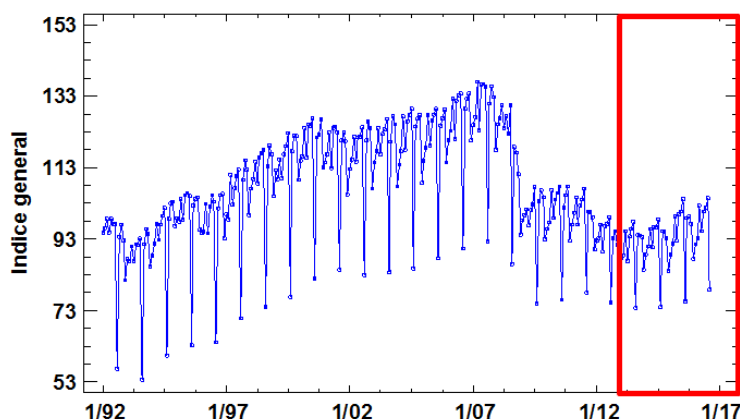


Figura 15 Quinta etapa 2014-2017

Tras el año 2014, en el que el crecimiento fue mínimo, en 2015 la industria comienza una recuperación en España. La producción industrial crece un 3,2%, con todos los sectores presentando tasas positivas. Y lo que es más importante: las empresas supervivientes han aprendido la importancia de la exportación, de la cualificación de los trabajadores y de la inversión en I+D+i. Incluso en las pequeñas y medianas empresas. En la actualidad, la industria es responsable de una proporción importante de los bienes exportados, con positivas consecuencias derivadas hacia otros sectores. Su empleo es más estable que en el resto de sectores.

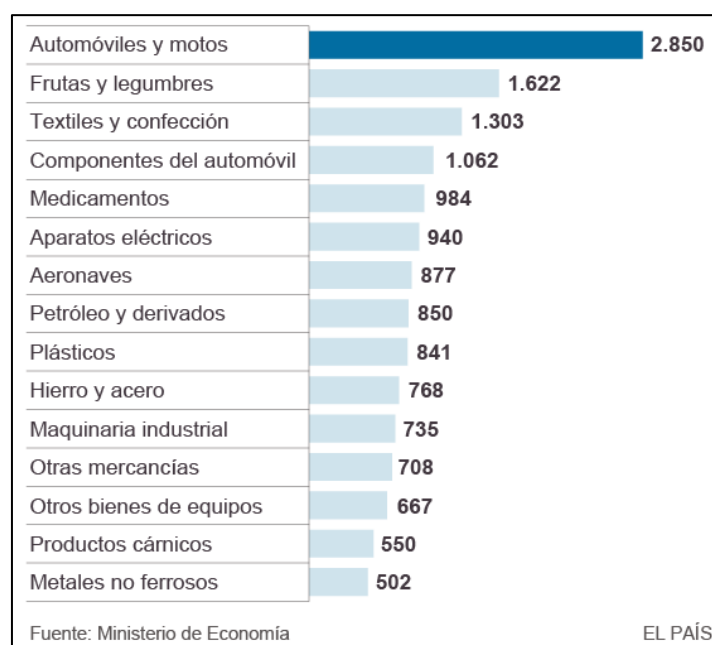


Figura 16 Exportaciones de España. 15 principales subsectores. En Millones de euros.

El esfuerzo se ha centrado en el diseño, la tecnología y la calidad. El sector del transporte (no sólo automóvil, sino también el material ferroviario o la construcción aeronáutica), el químico e incluso la alimentación han incorporado nuevas tecnologías. No obstante, deberán seguir buscándose soluciones a los problemas derivados del alto coste de la energía, que siguen estando presentes, la competencia procedente de los países emergentes y el cada vez más exigido respeto al medio ambiente.

Previamente, en 2014, España puso en marcha la Agenda para el Fortalecimiento del Sector Industrial que, a través de líneas de actuación y propuestas, permitirían mejorar la actividad

industrial y su adaptación a los nuevos retos globales, a los cambios tecnológicos y a la digitalización de la economía. La idea de reindustrialización de la economía está finalmente presente en las instituciones.

Las líneas de actuación son diez. Entre las de ámbito nacional destacan el estímulo de la demanda de bienes industriales y la mejora de la competitividad de los factores clave productivos. En el internacional, el apoyo a la internacionalización y diversificación de mercados de las empresas industriales. Las líneas de actuación transversales son incrementar la eficiencia, apoyar la profesionalización de las PYMES y fomentar la formación.

En la situación actual se dan las condiciones adecuadas para que la industria sea uno de los pilares que devuelvan a la economía española a una senda de crecimiento sostenible. Sin olvidar las debilidades de nuestro modelo industrial que obstaculizan la mejora de la eficiencia y la calidad de los procesos productivos. El reto obliga a fortalecer las bases de la competencia, ayudadas por el fomento de la investigación y la orientación hacia el exterior. La ayuda a las pequeñas y medianas empresas es otro requisito clave, así como la orientación comercial y productiva hacia nuevos mercados dinámicos y con mayor capacidad de compra.

3. Series temporales

3.1 Introducción

La finalidad del análisis de series temporales es estudiar cómo evoluciona una variable a lo largo de un periodo de tiempo. Las series pueden tener distinta periodicidad. Como ejemplo, la serie del Índice de Producción Industrial se toma con periodicidad mensual y la Demanda de Energía Eléctrica se puede obtener con periodicidad horaria. A partir de los datos mencionados será posible formular un modelo que represente la serie lo más fielmente posible con el fin de realizar previsiones.

3.2 Características de las series temporales

Una serie temporal se representa como z_t donde el subíndice t indica el periodo de observación. Se entiende que si se está en el periodo t , la previsión de la serie para $t+1$ será \hat{z}_{t+1} . Se denomina z_t a una serie temporal:

$$z_1, z_2, \dots, z_t, z_{t+1} \dots \quad (1)$$

Las características básicas de las series temporales son: tendencia, variabilidad y ciclo estacional.

La tendencia de una serie nos indica que la serie crece o decrece conforme pasa el tiempo, por lo que una serie sin tendencia es aquella que permanece estable en torno a un valor.

La variabilidad o volatilidad de la serie puede ser constante o puede cambiar, aumentando o disminuyendo a lo largo del tiempo. Si es constante se denomina homocedástica. Las series heterocedásticas van modificando su varianza con el tiempo.

El ciclo estacional se presenta en las series de periodicidad menor que la anual (mensual, semanal, etc.). Se pueden observar patrones de comportamiento cíclicos que producen el mencionado ciclo estacional de la serie que debería ser modelizado. Por ejemplo en la Figura 17 que ya se ha mostrado previamente en la Figura 3 se puede observar la existencia de ciclo y

tendencia. La tendencia se aprecia debido a que la serie de datos no permanece constante en torno a un valor y el ciclo se observa en la pauta repetida año tras año.

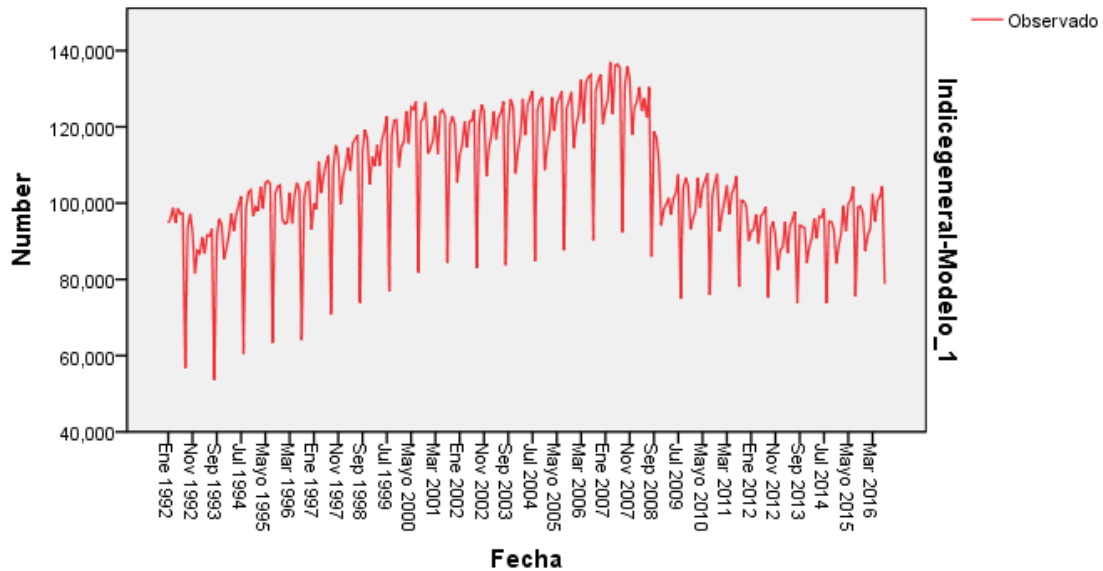


Figura 17 Serie del Índice de Producción Industrial

3.3 Estructura de dependencia

Para analizar la serie, hemos de estudiar la estructura que la genera. Esta estructura dependerá de cómo influyen en las observaciones del futuro aquellas del pasado. Para realizar este estudio se utilizan dos medios de información. La Función de Autocorrelación Simple (FAS) y la Función de Autocorrelación Parcial (FAP).

3.3.1 Función de Autocorrelación Simple (FAS)

Esta función proporciona la estructura de dependencia lineal de la serie y se compone de una sucesión de coeficientes:

$$\rho_1, \rho_2, \dots, \rho_k \quad (2)$$

Donde ρ_k será el coeficiente de autocorrelación entre una observación y otra situada k periodos más tarde. Los valores de los coeficientes estarán comprendidos entre $[-1, +1]$.

3.3.2 Función de Autocorrelación Parcial (FAP)

La Figura 18 muestra una posible estructura de dependencia de una serie temporal y se va a utilizar para explicar la diferencia entre la FAS y la FAP.

Así, por ejemplo, la observación z_5 recibe influencia de z_3 por dos vías, una directa entre z_3 y z_5 y otra indirecta de z_3 a z_4 y de z_4 a z_5 . La FAS proporciona como influencia de z_3 a z_5 la influencia total, es decir, las dos vías. La FAP es asimismo una sucesión de coeficientes de correlación pero solo obtenemos un coeficiente significativo donde hay relación directa. En el caso de la Figura 18 la FAP mostraría dos coeficientes significativos correspondientes al primero y segundo retardo.

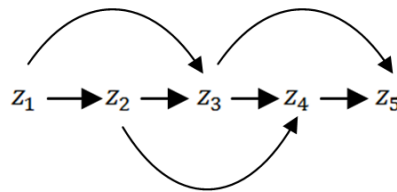


Figura 18 Estructura de dependencia de una serie temporal

3.4 Series estacionarias

3.4.1 Definición

Una serie es estacionaria si se cumplen las siguientes características:

- No debe tener tendencia
- Homocedástica
- Sin ciclo estacional
- La FAS depende solo del retardo k
- Influencia de las observaciones decreciente con el tiempo

Los modelos ARIMA se aplican a series estacionarias por lo éstas deben transformarse previamente para hacerlas estacionarias. Por ejemplo, para eliminar la tendencia de una serie z_t :

$$\nabla z_t = w_t = z_t - z_{t-1} \quad (3)$$

Siendo w_t la nueva serie y z_t las observaciones de la serie original y ∇z_t la serie diferenciada.

3.4.2 Operador de retardo

Dicho operador retarda la observación un periodo. Se define:

$$B \cdot z_t = z_{t-1} \quad (4)$$

Y por tanto:

$$B^k z_t = z_{t-k} \quad (5)$$

Aplicando el operador de retardo podemos obtener una nueva formulación del operador diferencia:

$$\nabla z_t = (1 - B)z_t = z_t - Bz_t = z_t - z_{t-1} \quad (6)$$

En el caso de que hubiese que tomar más de una diferencia para hacer estacionaria la serie la notación será la siguiente:

$$\nabla^n z_t = (1 - B)^n z_t \quad (7)$$

Siendo n el número de diferencias que se han de tomar.

3.5 Procesos autorregresivos AR

Los procesos autorregresivos se utilizan para representar la dependencia de los valores de una serie temporal a los de su pasado. En este caso la dependencia será lineal.

3.5.1 Proceso autorregresivo de primer orden AR(1)

Para que un proceso sea AR(1) ha de seguir la siguiente ecuación:

$$z_t = \phi z_{t-1} + a_t \quad (8)$$

$$a_t \sim N(0, \sigma^2)$$

$$E(a_t, a_{t+k}) = 0 \quad \forall k \neq 0$$

Con la siguiente estructura de dependencia:

$$z_1 \longrightarrow z_2 \longrightarrow z_3 \longrightarrow z_4 \longrightarrow z_5$$

Figura 19 Estructura de dependencia de una serie temporal

En la cual a_t es una perturbación aleatoria que se denomina ruido blanco y z_t es la observación t . El fin de introducir a_t es evitar que la serie sea determinista. Para ser estacionario el proceso AR(1) debe tener un parámetro $|\phi| < 1$.

3.5.2 Proceso autorregresivo de orden superior AR(p)

Los procesos autorregresivos de orden superior siguen el modelo:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (9)$$

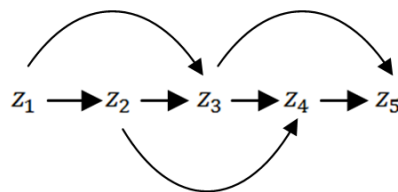


Figura 20 AR(p)

Esta ecuación muestra que la observación t estará afectada directamente por las p observaciones anteriores.

Pasaremos ahora a estudiar el caso del modelo AR(p) utilizando el operador de retardo:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = a_t \quad (10)$$

Denominamos $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$. Se llama a $\phi_p(B)$ polinomio característico, y sus raíces serán las que caractericen la forma de la FAS y la FAP. Utilizando la notación anterior el modelo AR(p) puede escribirse como:

$$\phi(B) z_t = a_t \quad (11)$$

La expresión obtenida es la general de un proceso autorregresivo.

La FAS de un proceso AR(p) depende de las raíces del polinomio característico. Será decreciente y cada raíz positiva aportará una estructura decreciente exponencial. Cada raíz negativa una estructura decreciente exponencial alternando coeficientes positivos y negativos. Finalmente cada pareja de raíces complejas conjugadas aportan una estructura sinusoidal decreciente y en este caso la serie presenta un ciclo no estacional de periodicidad aleatoria. En cualquier caso la FAS presenta bastante estructura decreciente. Este fenómeno se aprecia por ejemplo en la actividad solar.

La FAP puede demostrarse que presenta tantos polos significativos como orden tiene el proceso autorregresivo.

3.6 Procesos de Media Móvil MA

La estructura decreciente que se podía observar en la FAS de los procesos AR era debida a la larga memoria que tienen estos procesos. Es por esto por lo que los procesos AR tardan en absorber los impactos externos. Los procesos de media móvil tienen una memoria corta y modelizan situaciones de rápida absorción de las perturbaciones externas.

3.6.1 Procesos de media móvil de primer orden MA(1)

La ecuación de este tipo de procesos es:

$$z_t = a_t - \theta a_t \quad (12)$$

3.6.2 Procesos de media móvil de orden superior MA(q)

La característica principal de los procesos MA es la rápida absorción de impactos. Así, un MA(1) es capaz de absorber una alteración en una observación y un MA(q) en un número q de observaciones. Dichos procesos obedecerán a la siguiente ecuación:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}. \quad (13)$$

Aplicando el operador de retardo al modelo MA(q):

$$z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (14)$$

$$z_t = \theta_q(B) a_t \quad (15)$$

El proceso MA y el AR se comportan de una manera dual. La FAS y la FAP del MA son duales con respecto a los de los procesos AR, así: la FAS de un MA(q) tendrá q retardos significativos, la FAP de un MA(q) tendrá una estructura decreciente similar a la FAS de un AR.

3.7 Procesos ARMA(p,q)

Las series suelen ser una combinación de estructuras de media móvil y autorregresivas. Tendrán así una parte AR(p) y otra MA(q). La FAS y FAP de los procesos ARMA(p,q) son ambas decrecientes. Utilizando para el modelo ARMA(p,q) el operador de retardo la estructura polinómica es la siguiente:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (16)$$

$$\phi(B) z_t = \theta_q(B) a_t \quad (17)$$

3.8 Procesos Estacionales

3.8.1 Introducción

En la práctica se observan series entre las que hay una influencia directa una observación y la equivalente en el año posterior, en estos casos suele apreciarse un ciclo estacional como muestra la Figura 21.

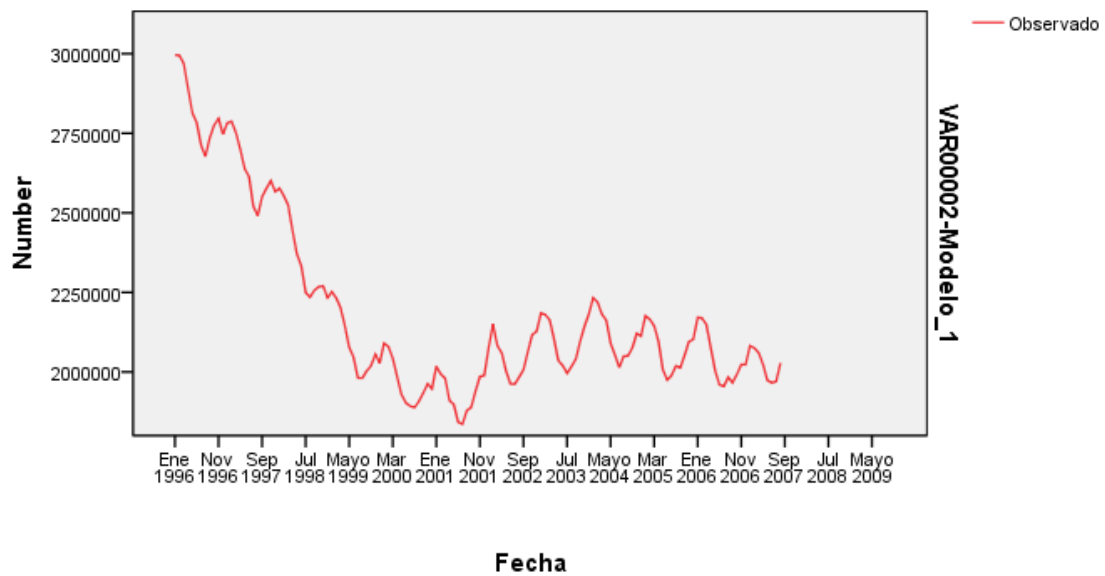


Figura 21 Ejemplo de serie con componente estacional

La serie será estacional cuando el valor esperado para cada periodo del año es constante, y diferente al de otros periodos del año:

$$E(z_t) = E(z_{t+s}) \quad (18)$$

Siendo la s el periodo de la estacionalidad. En la Figura 21 se puede observar cómo se repite una forma determinada cada año, en dicho caso la s sería 12. La Figura 22 muestra el esquema de relación estacional. Así F2017 recibe influencia directa de E2016 (en la estructura que llamamos regular) y también de F2016 a través del modelo estacional.

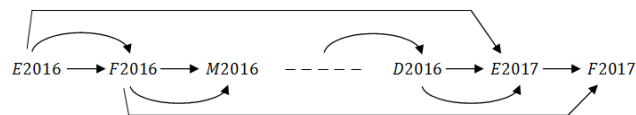


Figura 22 Modelo estacional con periodicidad anual ($s=12$)

3.8.2 Series estacionales no estacionarias

Al igual que para convertir una serie no estacionaria en una estacionaria se requería tomar una diferencia, para eliminar el ciclo estacional se han de tomar diferencias del orden estacional. El operador diferencia estacional de periodo s es definido como:

$$\nabla_s = 1 - B^s \quad (19)$$

La serie, sería definida de la siguiente manera:

$$\nabla_s z_t = (1 - B^s)z_t = z_t - z_{t-s} \quad (20)$$

3.8.3 Procesos $ARIMA(P, D, Q)_s$

En el caso de que una serie requiriese tomar una diferencia estacional y una regular se representaría de la siguiente manera:

$$w_t = \nabla \nabla_s z_t \quad (21)$$

Como se ha comentado una serie estacional podría tener un modelo AR, MA o un ARIMA. En la FAS y la FAP de estos modelos podremos observar la misma estructura que en los modelos no estacionales, pero separado por un número de retardos igual a la estacionalidad s . En el caso de una estacionalidad de 12 habría que fijarse en los retardos 12, 24...etc.

La ecuación que seguirán estos procesos $ARIMA(P, D, Q)_s$ será:

$$\Phi_P(B^S)\nabla_S^D z_t = \Theta_Q(B^S)a_t \quad (22)$$

Donde P son las diferencias tomadas para el AR, Q para el MA y D las diferencias estacionales.

3.9 Procesos generales ARIMA (p, d, q)x(P, D Q)_s

Sabiendo lo anterior se podría generalizar el modelo ARMA para series estacionarias añadiendo la dependencia regular y la dependencia estacional. De esta forma se obtendría el modelo ARIMA estacional multiplicativo. Que en su forma compacta es:

$$\Phi_P(B^S)\phi_p(B)\nabla^d \nabla_S^D z_t = \Theta_Q(B^S) \theta_q(B)a_t \quad (23)$$

En la cual $\Phi_P(B^S) = (1 - \Phi_1 B^S - \dots - \Phi_P B^{SP})$ es el operador AR estacional de orden P, $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ el operador AR de orden p, $\nabla^d = (1 - B)^d$ las diferencias regulares tomadas $\nabla_S^D = (1 - B^S)^D$ las diferencias estacionales, $\Theta_Q(B^S) = (1 - \Theta_1 B^S - \dots - \Theta_Q B^{SQ})$ el operador MA estacional de orden Q y, por último, $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ el operador MA de orden q. De esta forma el modelo $ARIMA(p, d, q)x(P, D, Q)_s$ general podría escribirse de la siguiente forma:

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)x(1 - \Phi_1 B^S - \dots - \Phi_P B^{SP})\nabla^d (\nabla_S^D) z_t \\ = (1 - \theta_1 B - \dots - \theta_q B^q)x(1 - \Theta_1 B^S - \dots - \Theta_Q B^{SQ})a_t \end{aligned} \quad (24)$$

Por ejemplo un modelo $ARIMA(1,1,1)x(1,1,1)_{12}$ tendrá la siguiente ecuación:

$$(1 - \phi_1 B)x(1 - \Phi_1 B^{12})\nabla\nabla_{12} z_t = (1 - \theta_1 B)x(1 - \Theta_1 B^{12})a_t \quad (25)$$

4. Metodología para el análisis de series temporales

La metodología para analizar las series temporales a utilizar será la Box-Jenkins, con el fin de que el ajuste y los pronósticos sean lo más correctos posible. Previamente a utilizar el método Box-Jenkins han de seguirse los siguientes pasos:

1. Toma de datos con un número de datos significativo.
2. Representación gráfica de la serie con el fin de comprobar si tiene tendencia o estacionalidad.
3. Comprobar homocedasticidad.
4. Eliminar la tendencia.

Para llevar a cabo el análisis han de seguirse las siguientes etapas:

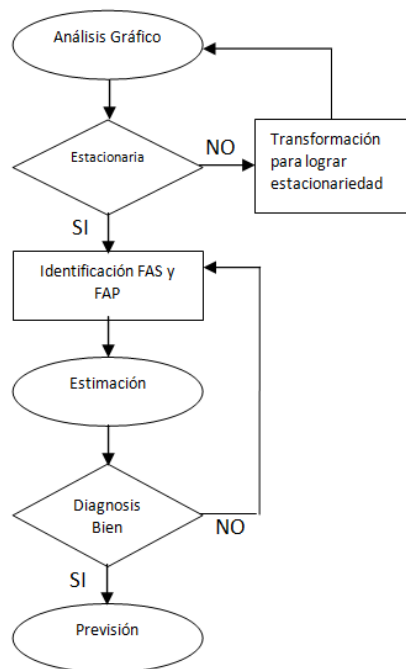


Figura 23 Etapas metodología Box-Jenkins

4.1 Identificación

Obtención de la estructura del proceso estocástico que ha generado los datos. De esta forma se obtienen los valores p , d y q del modelo ARIMA. Para ello se utilizan la FAS y la FAP.

	FAS	FAP
AR (p)	Muchos coeficientes no nulos que decrecen con el retardo como mezcla de exponenciales y senoides.	P primeros coeficientes no nulos y el resto 0.
MA (q)	q primeros coeficientes no nulos y el resto 0.	Muchos coeficientes no nulos que decrecen con el retardo como mezcla de exponenciales y senoides.
ARMA (p,q)	Decrecimiento hacia cero.	Decrecimiento hacia cero.

Tabla 1 Relación modelos con FAS y FAP. (Peña, 2005).

4.2 Estimación

En este paso se han de estimar los coeficientes utilizados en el modelo. Dicha estimación se lleva a cabo mediante algoritmos numéricos. En la realización del trabajo se han utilizado SPSS y R.

4.3 Comprobación de la Diagnosis

El objetivo de la etapa de verificación es comprobar si el modelo se ajusta suficientemente bien a los datos existentes. Esto es de gran interés puesto que podría decantar la elección de un modelo frente a otro, debido a un mejor ajuste. En la etapa de verificación se llevan a cabo tres comprobaciones, análisis de los coeficientes del modelo, bondad de ajuste y análisis de los residuos tras aplicar el modelo.

4.3.1 Análisis de los coeficientes del modelo

Se observan los valores del t estadístico de cada uno de los coeficientes del modelo con el fin de comprobar si su valor absoluto es mayor que 2. En caso de que así fuese nos indicaría la significancia individual de dicho coeficiente. Si el valor fuese menor que 2 sería un indicador de que el modelo no se está ajustando correctamente a los datos.

4.3.2 Bondad de ajuste

En este caso se comprueba el ajuste con el p -valor que se obtiene de los programas anteriormente mencionados. Este dato nos lo proporciona el test de Box-Pierce, puesto que dicho test indica problemas si el p -valor es menor que 0.05. A mayor p se tendrá mayor evidencia de que los residuos son ruido blanco.

4.3.3 Análisis de los residuos

El análisis se puede hacer mediante la observación de FAS y FAP tras aplicar a los datos el modelo planteado, dichas gráficas no deben tener ningún palo relevante por encima de las bandas de control. Un palo relevante es preocupante si está al comienzo de los retardos o en un múltiplo de la estacionalidad de los datos. Un ejemplo de esto se puede observar en la Figura 24. En ella se ve como los residuos son ruido blanco ya que el ruido blanco presenta FAS y FAP teóricamente cero y estimada dentro de las bandas de confianza.

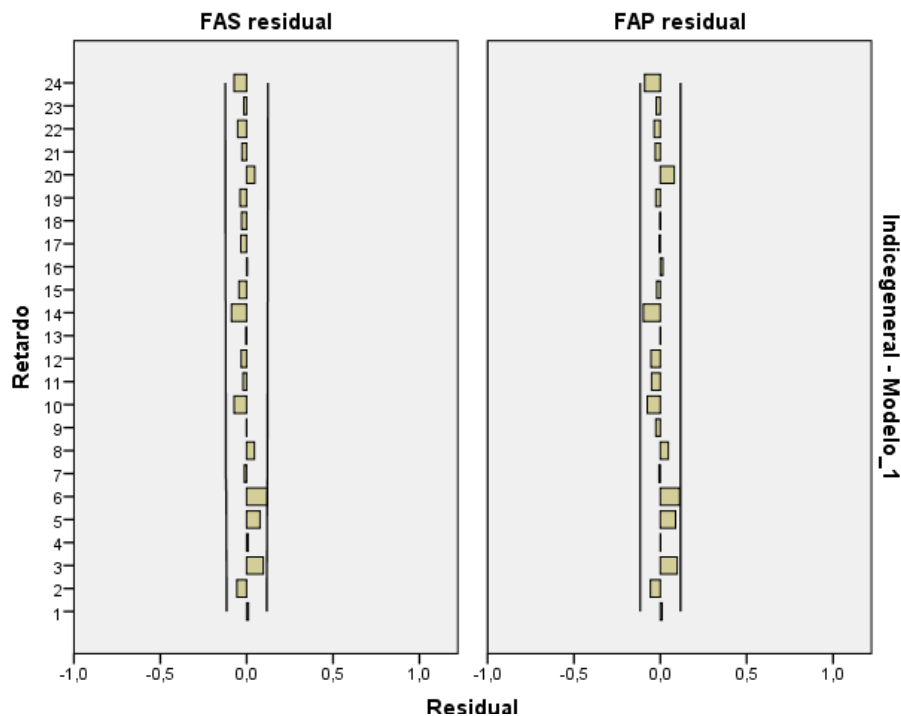


Figura 24 Verificación. Ruido blanco.

5. Análisis de intervención

Box y Tiao (1975) definieron el análisis de intervención. Este análisis introduce variables instrumentales (Impulso, escalón) para representar efectos determinísticos en series temporales. Es necesario introducir el análisis de intervención como paso previo a la definición de datos atípicos ya que éstos hacen uso del mismo.

5.1 Variable impulso

Partimos de un modelo que sigue una estructura ARIMA, que podemos escribir como:

$$\Phi_p(B^s)\phi_p(B)\nabla^d\nabla_s^D y_t = \Theta_q(B^s)\theta_q(B)a_t \quad (26)$$

De forma simplificada lo podemos escribir de la siguiente manera:

$$y_t = \psi(B) \cdot a_t \quad (27)$$

Supongamos que la serie está afectada por un evento externo en el momento $t = h$ conocido. Entonces, no se observa la serie y_t en todo momento, ya que en el instante h se observará una variable z_t relacionada por:

$$z_h = w_0 + y_h \quad (28)$$

Donde w_0 es el valor del impacto que ha afectado a la serie.

Debido a ese valor extraño no se pueden aplicar los modelos ARIMA estudiados anteriormente sin modificación pues se cometerán importantes errores en la estimación de los parámetros. Para explicar el modo en el que se resuelve este problema ha de definirse la variable impulso.

$$I_t^{(h)} = \begin{cases} 0 & t \neq h \\ 1 & t = h \end{cases} \quad (29)$$

Siguiendo z_t el modelo:

$$z_t = w_0 \cdot I_t^{(h)} + y_t \quad (30)$$

Si se utiliza la Ecuación (27):

$$z_t = w_0 \cdot I_t^{(h)} + \psi(B) \cdot a_t \quad (31)$$

Que es posible generalizar más aún si la serie recibe m impulsos:

$$z_t = \sum_{j=0}^m w_j \cdot I_t^{(h+j)} + \psi(B) \cdot a_t \quad (32)$$

Para comprender mejor esta ecuación y los modelos que vendrán después planteamos la siguiente igualdad:

$$BI_t^{(h)} = I_{t-1}^{(h)} = I_t^{(h+1)} \quad (33)$$

Esta igualdad es fácilmente explicable si se desarrollan los términos de la igualdad, para posteriormente extraer la regla general.

$$I_{t-1}^{(h)} = \begin{cases} 0 & t-1 \neq h \\ 1 & t-1 = h \end{cases} \Leftrightarrow I_t^{(h+1)} = \begin{cases} 0 & t \neq h+1 \\ 1 & t = h+1 \end{cases} \quad (34)$$

Puesto que se puede comprobar que:

$$t-j = h \Leftrightarrow t = h+j \quad (35)$$

Por lo tanto, la regla general será de la siguiente forma:

$$BI_t^{(h)} = I_{t-j}^{(h)} = I_t^{(h+j)} \quad (36)$$

De esta forma se obtiene una nomenclatura dependiente del factor B que se ha estudiado anteriormente:

$$z_t = w(B) \cdot I_t^{(h)} + \psi(B) \cdot a_t \quad (37)$$

Desarrollando el primer término de la suma:

$$w(B) = (w_0 + w_1 \cdot B + w_2 \cdot B^2 + \dots + w_m \cdot B^m) \quad (38)$$

$$w(B) \cdot I_t^{(h)} = (w_0 \cdot I_t^{(h)} + w_1 \cdot I_t^{(h+1)} + w_2 \cdot I_t^{(h+2)} + \dots + w_m \cdot I_t^{(h+m)}) \quad (39)$$

Dicho término $w(B)$ es un polinomio de orden m que tiene los coeficientes de intervención de cada uno de los m periodos siguientes. Se plantea ahora un modelo de comportamiento que indique el peso decreciente según se avanza en el tiempo:

$$z_t = \frac{w_0}{1 - \delta B} \cdot I_t^{(h)} + \psi(B) \cdot a_t \quad (40)$$

En este modelo $0 < \delta < 1$ para que los pesos sean decrecientes. La función de transferencia se puede reescribir, sabiendo que los pesos son decrecientes:

$$w_0(1 - \delta B)^{-1} = w_0(1 + \delta B + \delta^2 B^2 + \dots) = (w_0 + w_1 B + \dots + w_m B^m + \dots) \quad (41)$$

Donde $w_k = w_0 \delta^k$

Por todo lo anterior se puede escribir la función z_t como:

$$\begin{aligned} z_t &= y_t & , \quad t < h \\ z_{h+j} &= w_0 \delta^j + y_{h+j}, & j \geq 0 \end{aligned} \quad (42)$$

Como se puede ver en la ecuación (42), el efecto de un impacto (outlier) modifica las observaciones siguientes. Si no se tiene en cuenta, se cometerán importantes errores de estimación.

5.1.1 Variable escalón: ganancia

Las variables escalón son intervenciones con un efecto permanente sobre la serie, es decir, a diferencia de las variables impulso el efecto sobre la serie no es decreciente con el tiempo, sino que se mantiene constante. Aplicando los conceptos estudiados para la variable impulso se puede comprender mucho mejor este sistema de intervención. En este caso se utiliza una variable escalón.

$$S_t^{(h)} = \begin{cases} 0, & t < h \\ 1, & t \geq h \end{cases} \quad (43)$$

A partir de esta variable se puede representar el modelo de intervención, teniendo en cuenta que la serie sigue un modelo ARIMA:

$$z_t = w_0 \cdot S_t^{(h)} + \psi(B) \cdot a_t \quad (44)$$

De esta forma queda definida la serie de forma sencilla, puesto que se puede observar que dicha serie tendrá un comportamiento sin contaminar hasta el instante h , en el que la función escalón toma el valor 1, y de esta forma la serie queda modificada en un valor w_0 constante.

Por claridad se puede expresar de la siguiente manera:

$$\begin{aligned} z_t &= y_t & , \quad t < h \\ z_{h+j} &= w_0 + y_{h+j}, & j \geq 0 \end{aligned} \quad (45)$$

De esta forma se puede ver que la principal diferencia entre impulsos y escalones es que un impulso será transitorio mientras que un escalón es permanente. Aun así, una variable escalón podría verse como una serie de variables impulso, por lo que las similitudes son apreciables a simple vista.

5.2 Valores atípicos

Conociendo la teoría explicada anteriormente sobre el análisis de intervención, la comprensión de los valores atípicos se simplifica notablemente. La principal diferencia entre el análisis de intervención y los valores atípicos es que en el primero se conoce el momento temporal en el que estarán los efectos extraños sobre la serie, que en los ejemplos ha sido llamado h . Sin embargo, en la mayoría de los casos habrá diferentes alteraciones de la serie o errores que afectarán a la serie y que no son conocidos a priori. Es por esto por lo que el uso de valores atípicos será de gran importancia para el estudio de las series del sector industrial, puesto que si dichas alteraciones no son tenidas en cuenta esto podría afectar a las previsiones de las series, es decir, estarían sesgadas.

Para detectar esas anomalías en la serie de datos se establecerá un sistema para evaluar que datos han sido generados de forma diferente al resto. A continuación se estudiarán los atípicos más representativos junto con aquellos que serán de utilidad para el objetivo de este estudio.

5.2.1 Atípicos aditivos (AO)

Un atípico aditivo es aquel valor que en un instante h se genera de forma distinta al resto de la serie. Este atípico podría deberse, por ejemplo, a errores puntuales de medición. Para estudiarlo partiremos de la serie z_t , teniendo en cuenta que no conocemos el momento h en el que se dará el atípico mencionado. Dicha serie seguirá el modelo siguiente:

$$z_t = \begin{cases} y_t & t \neq h \\ y_t + w_A & t = h \end{cases} \quad (46)$$

Teniendo en cuenta, al igual que en análisis de intervención, que la serie y_t es la serie sin contaminar por los valores atípicos y que se ha supuesto que seguirá un modelo ARIMA.

$$y_t = \psi(B) \cdot a_t \quad (47)$$

Entonces el modelo que sigue la serie observable será:

$$z_t = w(B) \cdot I_t^{(h)} + \psi(B) \cdot a_t \quad (48)$$

Sabiendo que $I_t^{(h)} = 0, t \neq h$; $I_t^{(h)} = 1$, pero en este h será una incógnita puesto que no conocemos su valor. Utilizando las diferentes formas estudiadas de escribir el modelo ARIMA lo podemos reescribir de la siguiente forma:

$$\pi(B)y_t = a_t \quad (49)$$

Y por tanto,

$$\pi(B)(z_t - w_A \cdot I_t^{(h)}) = a_t \quad (50)$$

Estudiaremos ahora el efecto en los residuos de un atípico aditivo. Llamaremos e_t a los residuos que contienen el efecto del atípico aditivo a estudiar, frente a a_t , las innovaciones, que serán los residuos sin contaminar que por lo tanto tendrán forma de ruido blanco. Únicamente desarrollando la ecuación anterior obtenemos la relación entre residuos e innovaciones. Suponiendo un AR(1) partimos de que las observaciones han sido generadas de la forma:

$$e_t = z_t - \phi z_{t-1} \quad (51)$$

Aplicando esto a la ecuación mencionada:

$$(1 - \phi B)z_t = w_A(1 - \phi B)I_t^{(h)} + a_t \quad (52)$$

Que se puede escribir de la siguiente manera:

$$e_t = w_A \cdot x_t + a_t = w_A(I_t^{(h)} - \phi_1 I_t^{(h+1)} - \dots - \phi_p I_t^{(h+p)}) + a_t \quad (53)$$

Por lo que todos los p residuos posteriores a h estarán afectados:

$$e_{h+j} = a_{h+j} - \phi_j w_A \quad j \geq 0 \quad (54)$$

Conociendo $\pi(B)$, planteamos por último la ecuación general.

$$e_{h+j} = a_{h+j} - \pi_j w_A \quad j \geq 0 \quad (55)$$

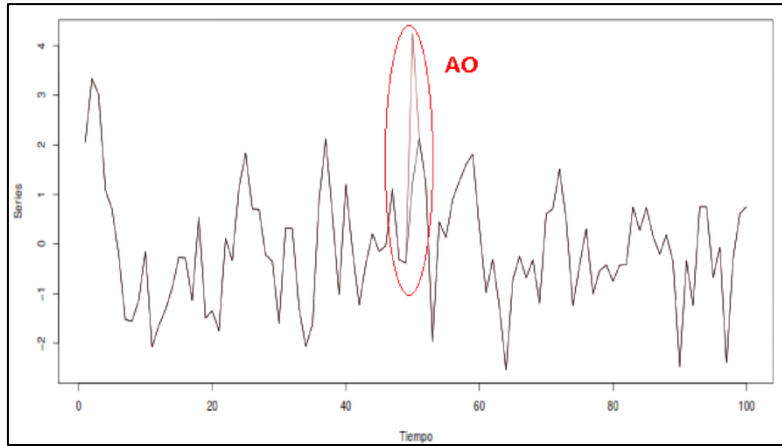


Figura 25 Atípico aditivo modelo AR (Galeano, 2011)

5.2.2 Innovativo (IO)

Un atípico innovativo es un cambio en la innovación de la serie temporal, la cual contiene el efecto no sistemático que las distintas variable ejercen influencia sobre la serie. Dicho efecto está definido de la siguiente forma:

$$z_t = y_t + \psi(B)w_I I_t^{(h)} \quad (56)$$

Donde y_t es la serie ARIMA original sin contaminar, z_t es la serie actual, $\psi(B)$ es el modelo MA y w es el valor del atípico. Entre y_t y z_t existe la siguiente relación:

$$z_t = \begin{cases} y_t, & t < h \\ y_t + w_I \psi_j, & t = h + j, \quad j \geq 0 \end{cases} \quad (57)$$

Por lo tanto cuando la $t = h$ la función $I_t^{(h)}$ toma el valor 1 y por lo tanto la serie se ve afectada por el atípico innovativo.

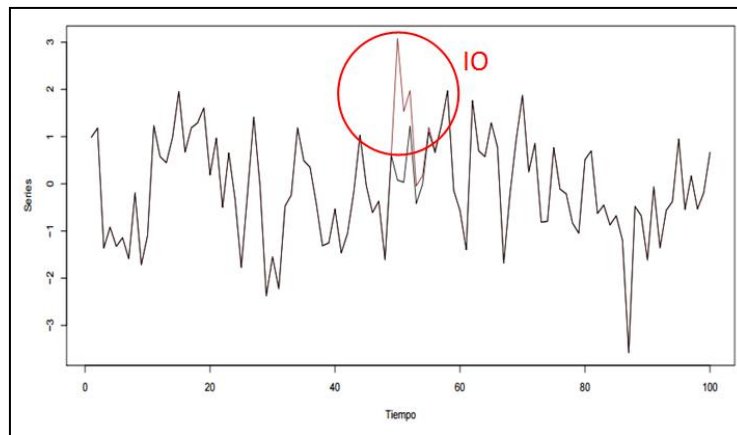


Figura 26 Atípico innovativo modelo ARIMA, (Galeano, 2011)

5.2.3 Cambio de nivel (LS)

Otro de los efectos más comunes que pueden darse en una serie temporal es el cambio de nivel. Dicho efecto se puede entender como una serie de atípicos aditivos que comienzan a afectar a la serie en un momento h , sin embargo, a diferencia de los atípicos aditivos en el cambio de nivel el efecto permanece constante. Para explicar el efecto de forma más sencilla se utilizará la teoría explicada en la variable escalón, por la cual la serie sufre un cambio de nivel en un instante h . El modelo para el cambio de nivel es el siguiente.

$$z_t = w_0 \cdot S_t^{(h)} + \psi(B) \cdot a_t \quad (58)$$

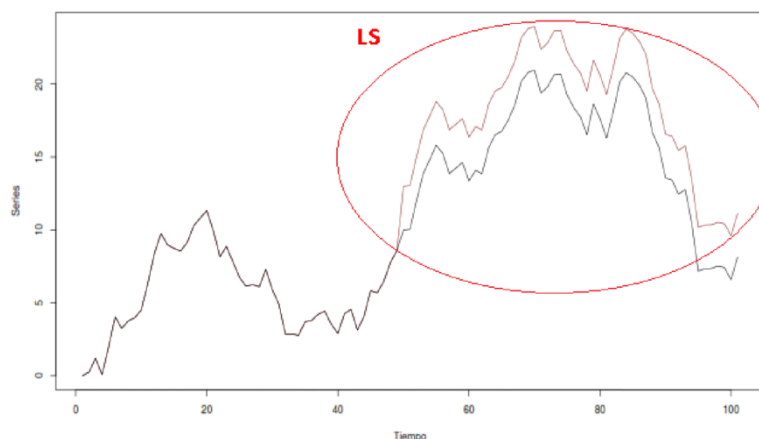


Figura 27 Cambio de nivel modelo ARIMA, (Galeano, 2011)

6. Análisis de las series temporales del sector industrial

6.1 Introducción

Para analizar el sector industrial se utilizarán las series más representativas del mismo. Se analizará también el paro mensual debido a la interrelación de este con la situación económica. La fuente de todos los datos utilizados en las series es el Instituto Nacional de Estadística.

6.2 Índice de Producción Industrial (IPI)

6.2.1 Definición

El Índice de Producción Industrial (IPI) es un indicador de coyuntura que mide cómo evoluciona mensualmente la actividad de producción del sector industrial. Dentro de este índice con alto grado de agregación se encuentran según su destino económico los siguientes sectores: bienes de consumo, energía, bienes de equipo y bienes intermedios. Dicho indicador mide los cambios en la producción industrial sin que afecte la variación de los precios.

6.2.2 Datos

Los datos utilizados comienzan en Enero de 1992 y se toman con una periodicidad mensual hasta el mes actual. La serie del Índice de Producción Industrial frente al tiempo es la siguiente.

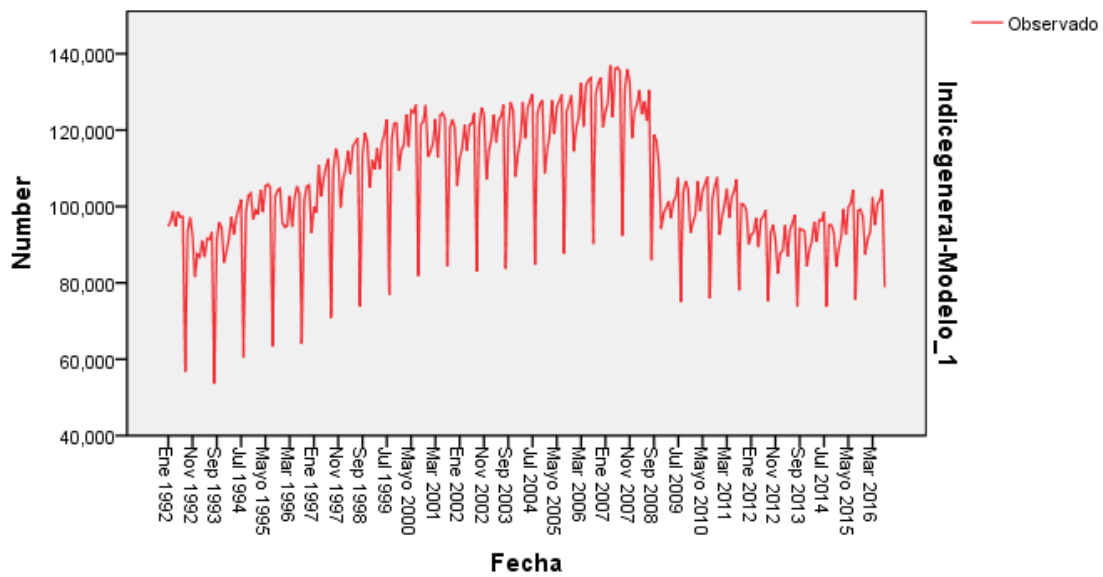


Figura 28 Gráfica Índice de Producción Industrial

En dicha gráfica se puede observar claramente una tendencia que es ascendente desde el comienzo de la serie temporal hasta finales del año 2007, donde se encuentra una caída que podemos suponer se debe al efecto de la crisis económica. Posteriormente y hacia el final de la serie se puede observar una leve tendencia al alza. También es llamativo el descenso anual en los meses de Agosto, que se debe a las vacaciones estivales.

Para poder comprender mejor el comportamiento de dicha serie temporal hemos de utilizar las herramientas mencionadas en la parte teórica inicial, para ello comenzaremos observando las funciones de autocorrelación simple y parcial, que en este Índice son las siguientes:

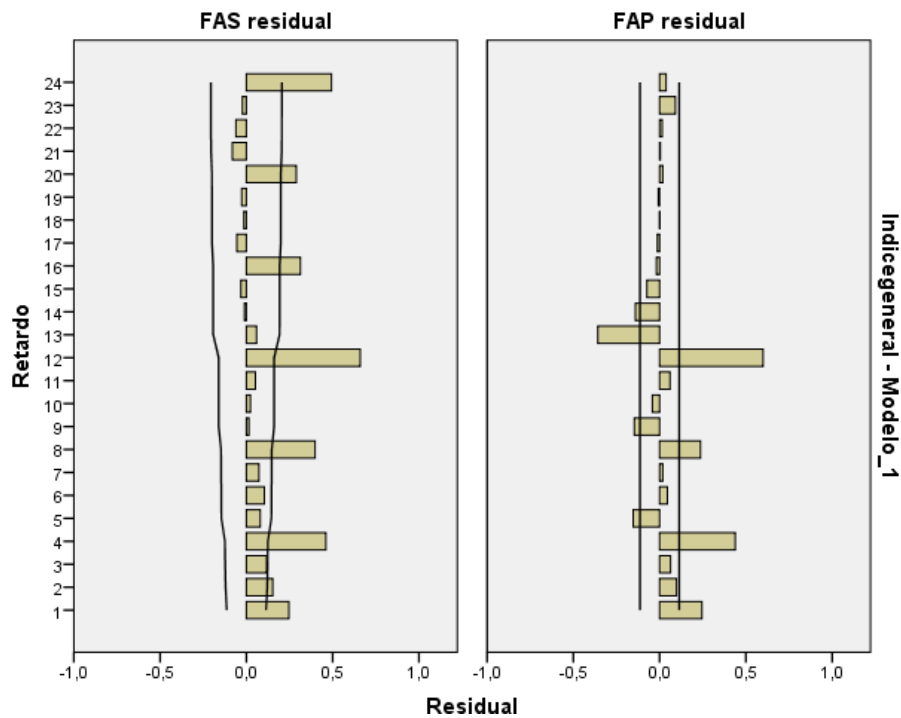


Figura 29 FAS y FAP del IPI previos al análisis

6.2.3 Análisis de datos

Debido a que el IPI es de gran importancia en el trabajo y por ser la primera serie analizada se documentará todo el proceso de ajuste hasta obtener una serie estacionaria. Partiremos de los datos de la serie que como habíamos visto antes tienen la siguiente gráfica:

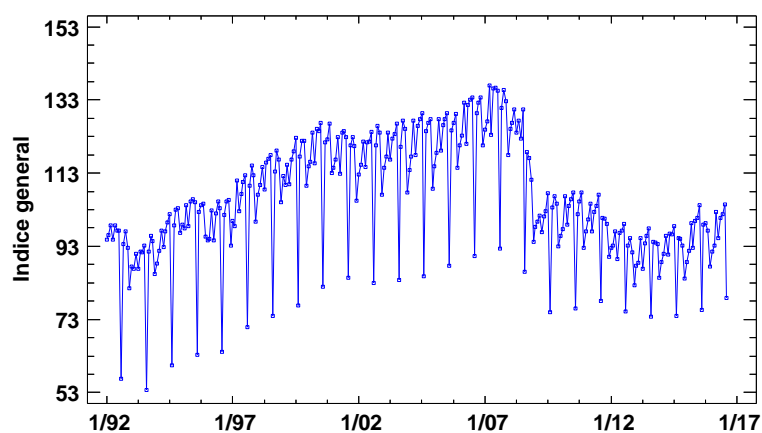


Figura 30 Gráfica del IPI

Como habíamos observado anteriormente la serie del IPI no es estacionaria debido a la tendencia y al ciclo estacional. Para eliminar el ciclo tomamos una diferencia estacional.

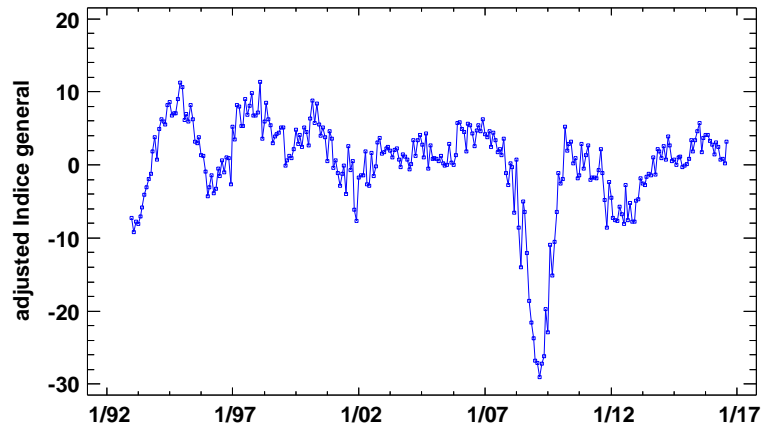


Figura 31 Gráfica IPI tras tomar una diferencia estacional

A pesar de haber tomado una diferencia estacional se observa que la serie sigue teniendo tendencia, de forma que será necesario tomar una diferencia regular para tratar de quitar dicha tendencia. Tomando ambas diferencias $\nabla \nabla_{12} IPI$ el resultado obtenido es el siguiente:

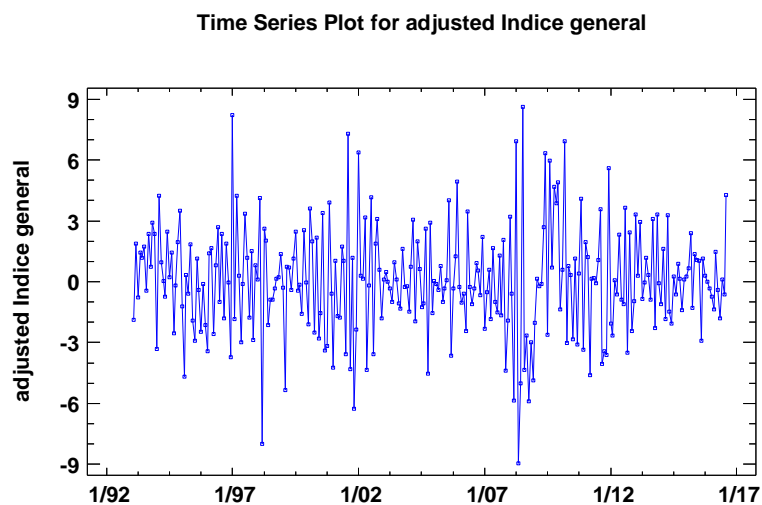


Figura 32 Gráfica IPI tras tomar una diferencia estacional y una regular

A partir de la FAS y la FAP obtenidas y sabiendo que la serie es ya estacionaria utilizamos un modelo ARIMA para obtener la estructura de los datos. Aplicamos una estructura ARIMA(1,1,2)(2,1,0)₁₂ y escribimos la ecuación que rige su comportamiento:

$$(1 - \phi_1 B) \cdot (1 - \Phi_1 B^{12} - \Phi_2 B^{24}) \nabla \nabla_{12} z_t = (1 - \theta_1 B - \theta_2 B^2) \cdot a_t \quad (59)$$

Los residuos obtenidos al aplicar la estructura enunciada a la serie de datos del índice de producción industrial son los siguientes:

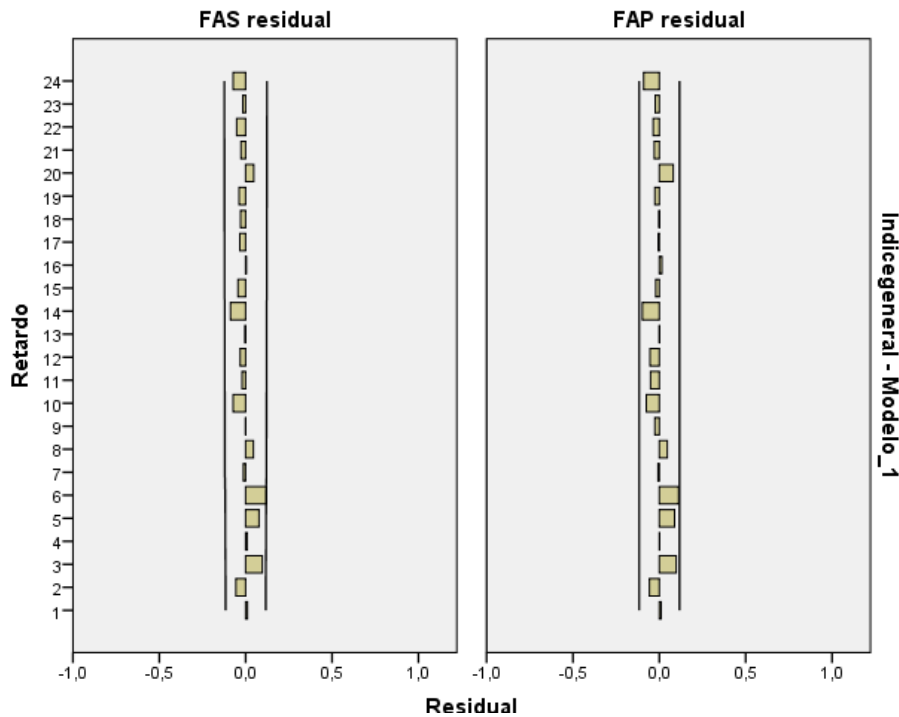


Figura 33 FAS y FAP tras aplicar el modelo ARIMA(1,1,2)(2,1,0)₁₂

En los resultados se puede observar que no hay ninguna autocorrelación superando las bandas de confianza, por lo que los residuos del modelo son ruido blanco.

La ecuación del modelo estimado será:

$$(1 + 0,967B) \cdot (1 + 0,291B^{12} + 0,245B^{24}) \nabla \nabla_{12} z_t = (1 + 0,483B - 0,459B^2) \cdot a_t - 0,03 \quad (60)$$

$(-5,913) \quad (-5,031) \quad (-4,256) \quad (-2,694) \quad (4,265) \quad (0,624)$

Donde los números entre paréntesis son los estadísticos t. Un parámetro es significativo cuando $|t| > 2$. Además, con el fin de comprobar que los residuos obtenidos son ruido blanco comprobamos el resultado del test Box-Pierce, que debe presentar un p-valor mayor que 0,05.

En nuestro caso:

$$p - \text{valor} = 0.260 \quad (61)$$

Por lo que no podemos rechazar la hipótesis nula de que los residuos son ruido blanco. A continuación vamos a estimar los posibles outliers que haya dentro de esta serie.

6.2.4 Outliers

Valores atípicos

			Estimación	SE	t	Sig.
Indice general-Modelo_1	Feb 2008	Tendencia local	-2,417	,341	-7,087	,000
	Jun 2008	Aditivo	-5,709	1,489	-3,834	,000
	Abr 2009	Tendencia local	1,864	,354	5,272	,000
	Ago 2009	Innovador	10,867	2,262	4,805	,000

Tabla 2 Outliers IPI

Como puede observarse en la Tabla 2 los Outliers se encuentran en el entorno de la crisis del 2007 con unas bajadas iniciales que se compensan con unas subidas en el año siguiente.

6.3 Número total de parados

6.3.1 Definición

Aunque la serie del paro no es específica del sector industrial se ha introducido en este trabajo por su relación evidente con el sector industrial. Vamos a estudiar la serie de paro registrado para todos los parados, tanto hombres como mujeres.

6.3.2 Datos

Los datos de la serie del paro total comienzan en Enero de 1996 y tienen una periodicidad mensual hasta Mayo de 2017. La serie obtenida es la siguiente:

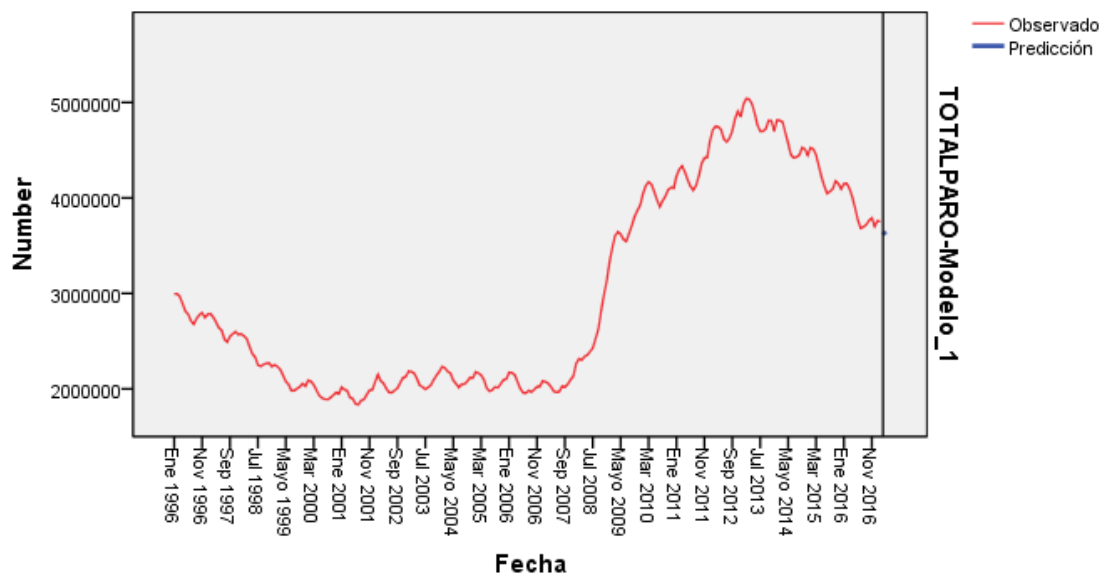


Figura 34 Gráfica Paro Total

En la gráfica del paro es muy llamativo el crecimiento casi exponencial que se da durante el año 2008 pasando de tener cerca de 2.000.000 de parados en el año 2007 a tener 3.000.000 tan solo un año después. Es esperable que dicho comportamiento esté recogido en los Outliers que se obtendrán tras el análisis de la serie. Al igual que en la anterior serie comenzamos observando la FAS y la FAP.

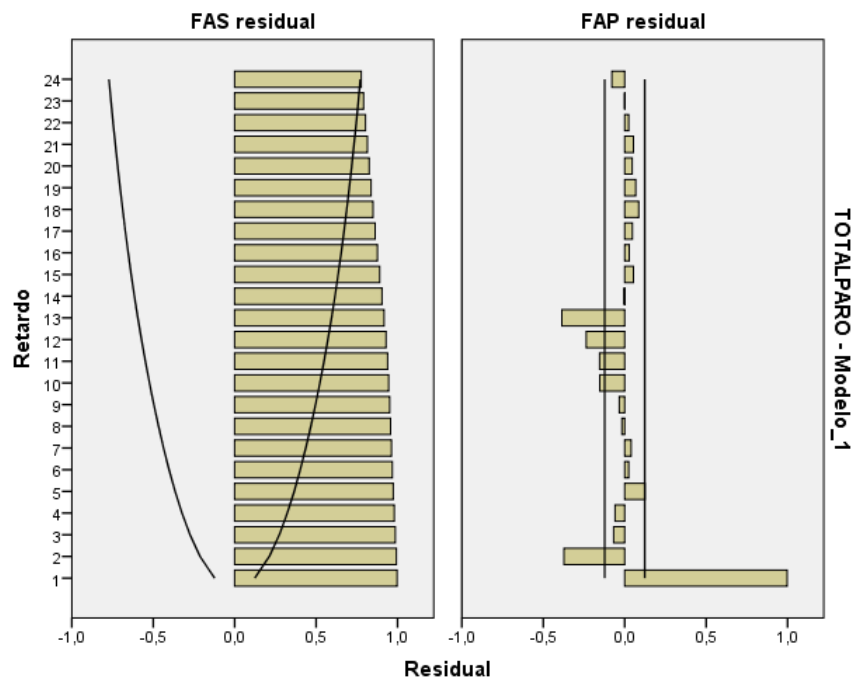


Figura 35 FAS y FAP de Paro total

6.3.3 Análisis de datos

Para quitarle la tendencia a la serie habrán de tomarse dos diferencias regulares. Además habrá de tomarse una diferencia estacional para quitar el ciclo. Tomando únicamente una diferencia regular y una estacional la serie quedaría tal como se ve en la Figura 36. Como puede observarse la serie no está estabilizada por lo que es necesario tomar una segunda diferencia regular. El resultado se muestra en la Figura 37.

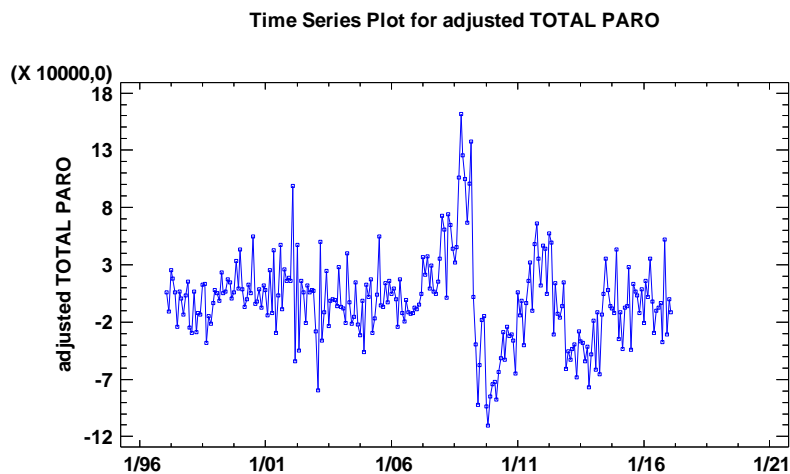


Figura 36 Gráfica Paro total tras quitar una diferencia regular y una estacional

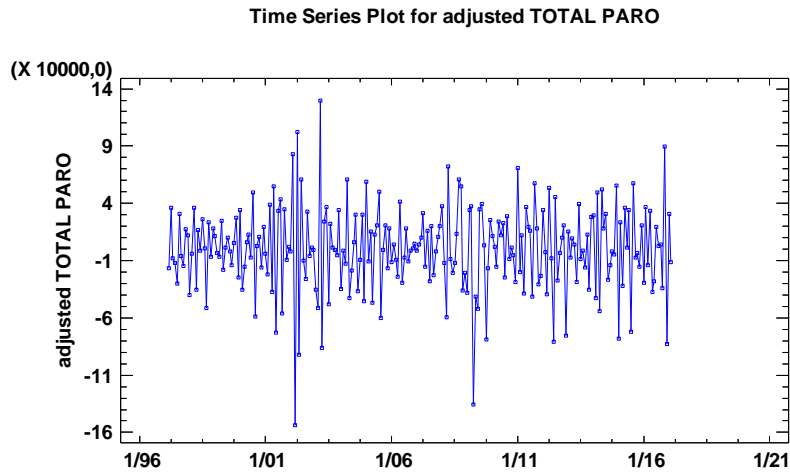


Figura 37 Gráfica Paro total tras quitar dos diferencias regulares y una estacional

Se ajusta un modelo ARIMA(0,2,1)(0,1,1)₁₂.

$$\nabla^2 \nabla_{12} z_t = (1 - \theta_1 B) \cdot (1 - \Theta_1 B^{12}) a_t \quad (62)$$

$$\nabla^2 \nabla_{12} z_t = (1 - 0,694B) \cdot (1 - 0,372B^{12}) a_t + 68,308 \quad (63)$$

(15,930)(6,190) (0,219)

La Figura 38 muestra la FAS y FAP de los residuos y el p-valor del test de Box-Pierce es 0,471 por lo que no podemos rechazar la hipótesis nula de que los residuos son ruido blanco y consecuentemente el modelo es válido.

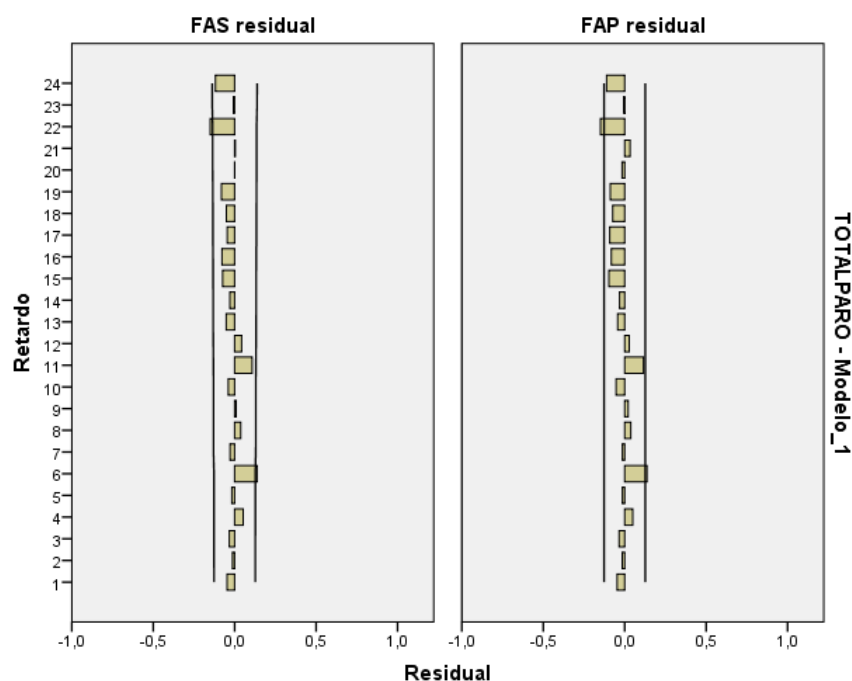


Figura 38 FAS y FAP tras aplicar un modelo $ARIMA(0,2,1)(0,1,1)_{12}$

6.3.4 Outliers

			Estimación	SE	t	Sig.
TOTAL PARO-Modelo_1	Feb 2002	Aditivo	66833,222	12158,896	5,497	,000
	Abr 2008	Tendencia local	74465,141	15040,872	4,951	,000
	Oct 2008	Tendencia local	53321,869	14608,595	3,650	,000
	Jun 2009	Innovador	-112415,289	25221,540	-4,457	,000

Tabla 3 Outliers Paro total

Los Outliers obtenidos para este modelo se encuentran en la

Tabla 3.

6.4 Índices de Precios Industriales (IPRI)

6.4.1 Definición

El objetivo del Índice de Precios Industriales es medir cómo evoluciona en el tiempo el precio medio de los productos del sector industrial que son vendidos en el mercado nacional en su primera etapa de comercialización sin tener en cuenta diversos gastos como impuestos, transporte y comercialización. En este índice se excluye el sector de la construcción.

6.4.2 Datos

Los datos comienzan en Enero de 1991 y terminan en y dichos datos son mensuales. La gráfica obtenida es la siguiente:

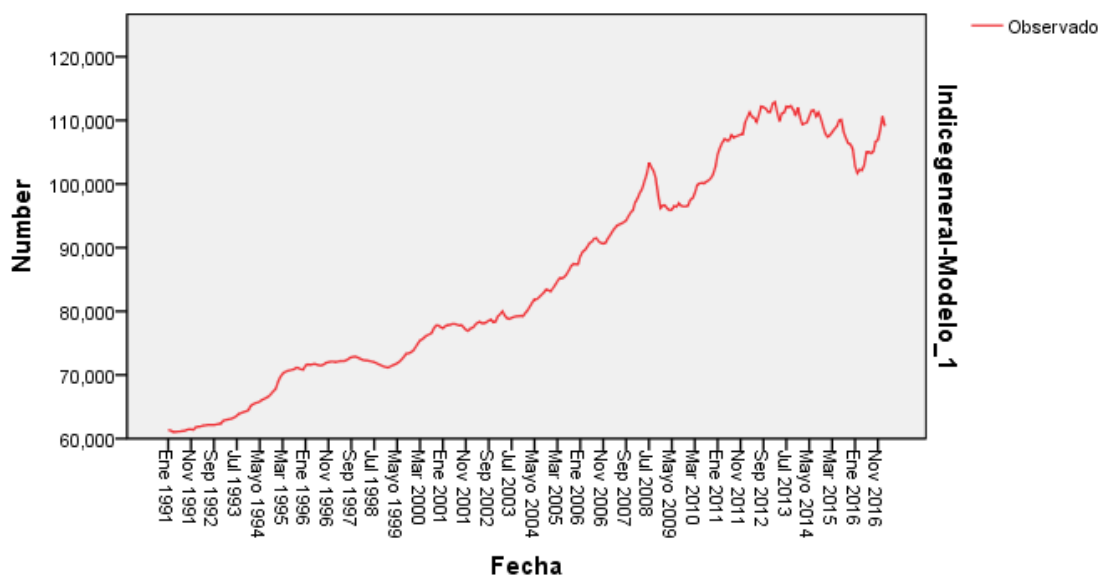


Figura 39 Gráfica Índice de Precios Industriales

En la gráfica se puede observar una tendencia ascendente hasta el año 2008, donde hay un descenso brusco con una etapa posterior en la que los precios se estabilizan entre 2011 y 2015 con una caída hasta 2016. En 2017 los precios remontan, en secciones posteriores se comprobará mediante la predicción de los modelos ARIMA con outliers si los modelos afirman que la tendencia seguirá al alza. En los precios no se observa una variación según el mes que se observaba en series anteriores.

Obtenemos la FAS y la FAP de la serie inicial.

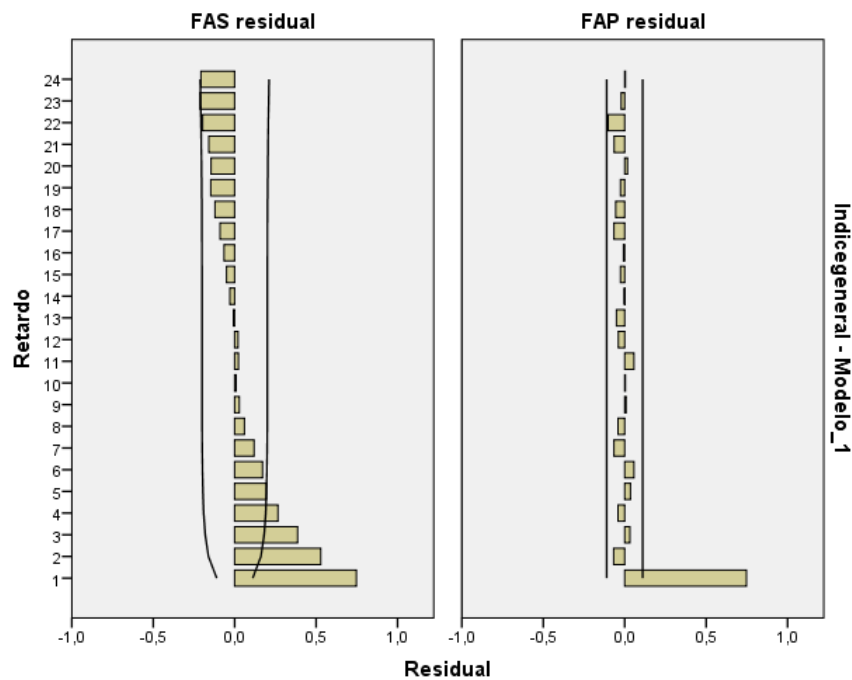


Figura 40 FAS y FAP datos IPRI

6.4.3 Análisis de Datos

Puesto que como habíamos observado anteriormente la serie de datos no era lineal pero sin embargo sí que era estacional se han de aplicar cambios para que sea estacionaria. Por ello se toma una diferencia regular, con el fin de quitarle la tendencia y una estacional para quitarle la estacionalidad.

Sabiendo que los datos están estacionarizados podemos, observando la FAS y la FAP obtenidas utilizamos un modelo ARIMA para obtener la estructura de los datos. Aplicamos una estructura ARIMA(1,1,0)(1,1,1)₁₂ y escribimos la ecuación que rige su comportamiento:

$$(1 - \phi_1 B) \cdot (1 - \Phi_1 B^{12}) \nabla \nabla_{12} z_t = (1 - \Theta_1 B^{12}) a_t \quad (64)$$

Aplicando el modelo propuesto obtenemos los siguientes residuos:

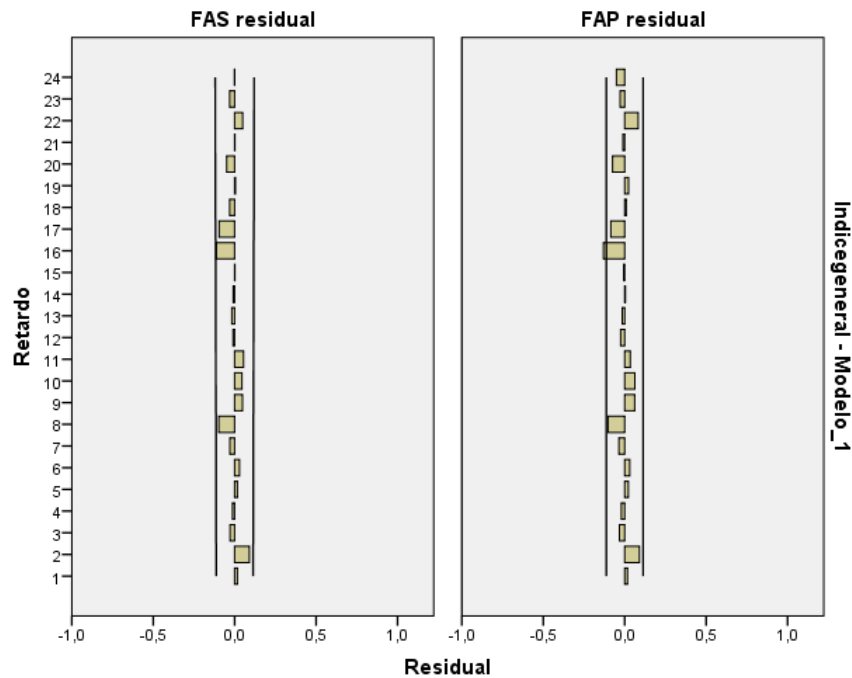


Figura 41 FAS y FAP del IPRI tras aplicar el modelo ARIMA(1,1,0)(1,1,1)₁₂

Al igual que en los modelos analizados anteriormente se observa que ninguno de los coeficientes de correlación de los retardos supera las bandas de confianza de la FAS y de la FAP. Esto indica que los residuos son ruido blanco, implicando que la estructura de los datos ha sido capturada en el modelo ARIMA planteado teniendo en cuenta la presencia de Outliers que se analizará posteriormente. Para tener mayor certeza de que la estructura es correcta se han de observar los parámetros del modelo en la ecuación (65), que como puede verse son mayores en valor absoluto que 2, por lo que dichos parámetros son significativos.

$$\begin{matrix} (1 - 0,455B) \cdot (1 - 0,184B^{12}) \nabla \nabla_{12} z_t = (1 - 0,971B^{12}) a_t \\ (11,160) \quad \quad (2,876) \quad \quad \quad (14,316) \end{matrix} \quad (65)$$

De nuevo comprobamos además el resultado del test Box-Pierce, que en este caso es:

$$p - \text{valor} = 0.371 \quad (66)$$

El p-valor es mayor que 0,05 por lo que no podemos rechazar la hipótesis nula de que los residuos son ruido blanco.

6.4.4 Outliers

Valores atípicos

			Estimación	SE	t	Sig.
Indice general-Modelo_1	Jul 2008	Aditivo	,054	,010	5,416	,000
	Nov 2008	Innovador	-,096	,019	-5,050	,000
	Mar 2013	Innovador	-,098	,019	-5,233	,000
	Abr 2013	Aditivo	-,043	,010	-4,237	,000
	Dic 2013	Aditivo	,090	,010	8,675	,000
	Feb 2014	Cambio de nivel	-,067	,017	-3,923	,000
	Ago 2014	Aditivo	-,041	,010	-4,027	,000
	Ago 2015	Innovador	-,097	,019	-5,132	,000
	Ene 2016	Innovador	-,146	,019	-7,551	,000
	Jun 2016	Cambio de nivel	,078	,017	4,531	,000
	Oct 2016	Innovador	,076	,019	4,052	,000
	Dic 2016	Innovador	,084	,019	4,473	,000
	Ene 2017	Aditivo	,072	,011	6,594	,000

Tabla 4 Outliers IPRI

Se observa que hay gran cantidad de valores atípicos del IPRI, esto se puede deber a la gran cantidad de factores que afectan a dicho índice.

6.5 Índices de Precios de Exportación de Productos Industriales

6.5.1 Definición

El Índice de precios de exportación (IPRIX) mide como cambian los precios de los productos industriales mensualmente. Los productos estudiados son aquellos que se fabrican en el mercado interior y que son vendidos posteriormente en el mercado exterior.

6.5.2 Datos

Los datos comienzan en Enero de 2005 y dichos datos son mensuales. La gráfica obtenida es la siguiente:

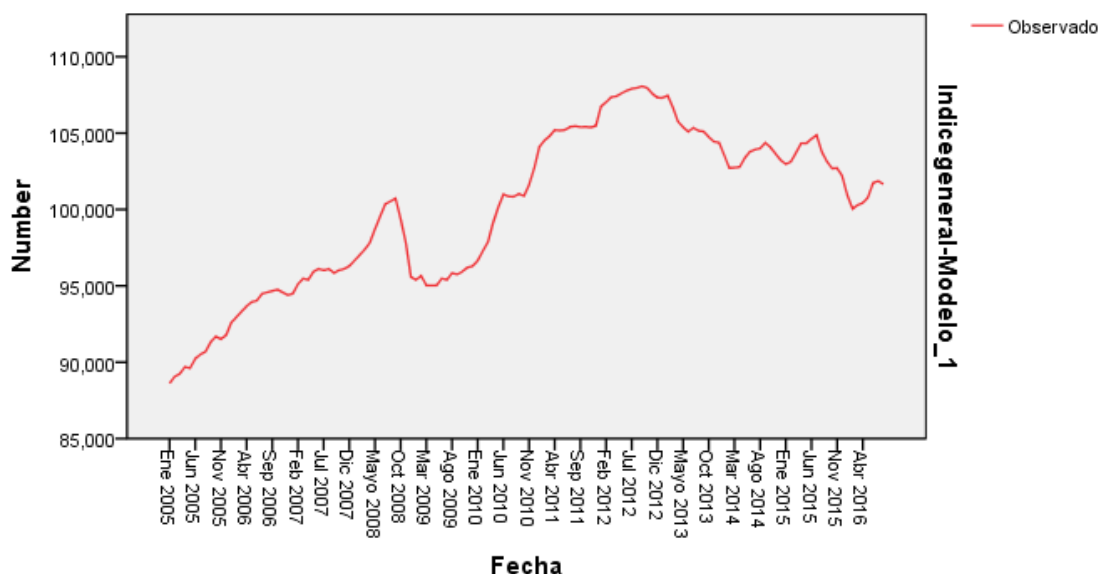


Figura 42 Gráfica del Índice de precios de exportación (IPRIX)

En la gráfica de los precios de Exportación se puede ver una tendencia ascendente hasta el año 2008, año de la quiebra de Lehman Brothers, en el que el índice cae notablemente. Sin embargo se ve una cierta recuperación posterior y un final en el que se estanca. De nuevo utilizaremos la FAS y la FAP para poder obtener la estructura de la serie. Dichas funciones son las siguientes:

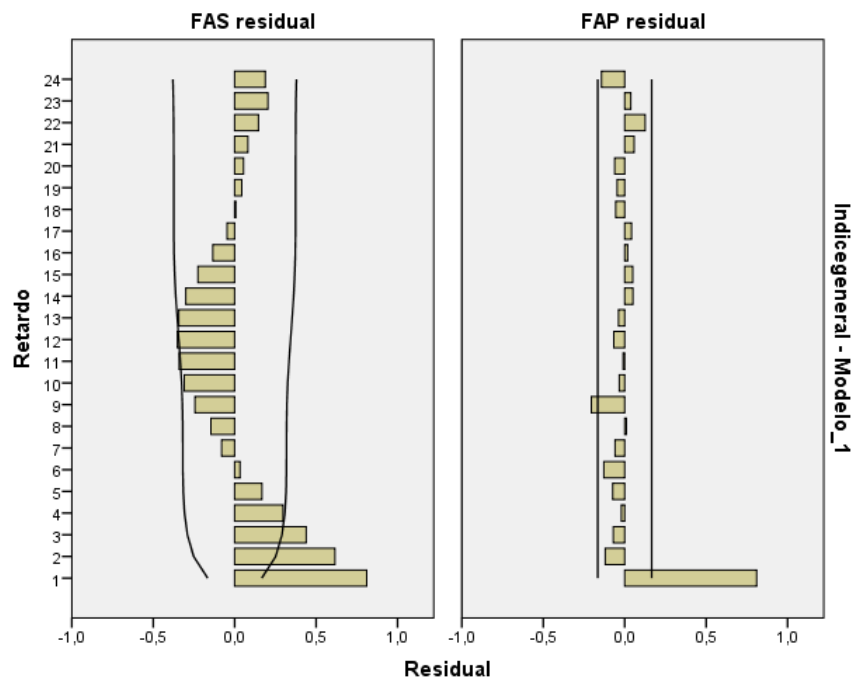


Figura 43 FAS y FAP del IPRIX

6.5.3 Análisis de datos

En la gráfica de la serie observamos que no es estacionaria debido a que presenta tendencia. Con el fin de eliminar la tendencia de la serie hemos de tomar una diferencia regular. Habiendo obtenido los resultados anteriores podemos observar que existen en la FAS y en la FAP un palo en el retardo 1 que supera las bandas de confianza. Al observar que tiene más estructura el de la FAS se planteará la hipótesis de que los datos siguen una estructura ARIMA(1,1,0). Por lo tanto la ecuación que rige su comportamiento teniendo en cuenta la diferencia ya tomada es la siguiente:

$$(1 - \phi_1 B) \cdot \nabla z_t = a_t \quad (67)$$

Si aplicamos el modelo planteado a la serie de datos obtenemos los siguientes residuos:

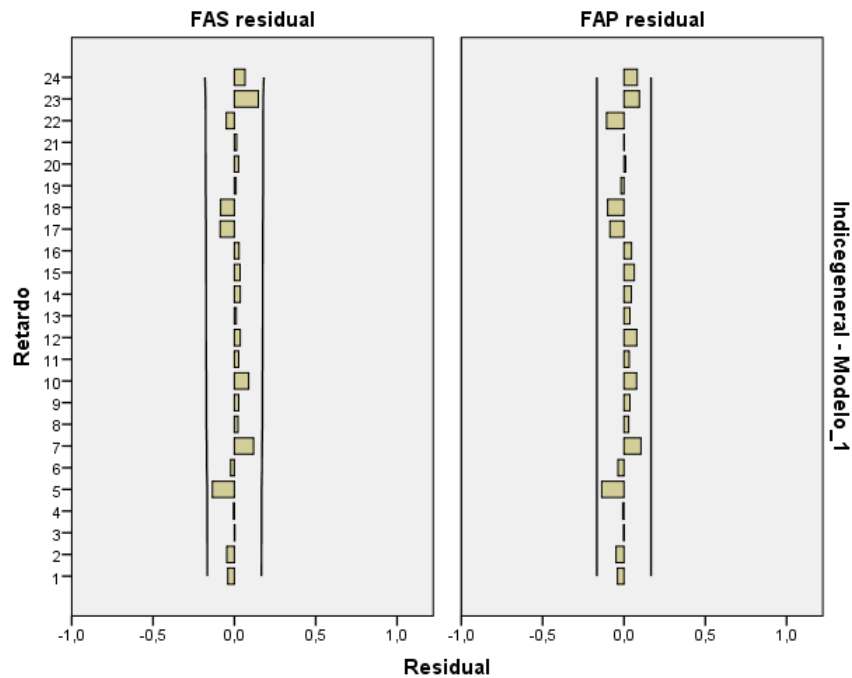


Figura 44 FAS y FAP tras aplicar un modelo ARIMA(1,1,0)

Puesto que ningún palo supera los límites de control podemos afirmar que lo obtenido es ruido blanco. Sin embargo al igual que en el resto de series para tener una mayor certeza de que el modelo es correcto hemos de comprobar los Parámetros del modelo ARIMA arrojados por el programa y el p-valor. En este caso los valores obtenidos son los siguientes:

$$\begin{matrix} (1 - 0,59B) \cdot \nabla z_t = a_t + 0,146 \\ (9,132) \qquad \qquad \qquad (1,972) \end{matrix} \quad (68)$$

$$p - valor = 0.904 \quad (69)$$

Puesto que en valor absoluto el valor del parámetro t es mayor que 2 podemos afirmar que el modelo ARIMA(1,1,0) es correcto, al igual que el p-valor, puesto que al ser mucho mayor que 0,05 por lo que no podemos rechazar la hipótesis nula de que los residuos son ruido blanco.

6.5.4 Outliers

Valores atípicos

			Estimación	SE	t	Sig.
Indice general-Modelo_1	Oct 2008	Innovador	-1,522	,358	-4,251	,000
	Dic 2008	Cambio de nivel	-1,527	,312	-4,891	,000
	Ene 2012	Cambio de nivel	1,091	,307	3,551	,001
	Ago 2015	Innovador	-1,280	,358	-3,574	,000

Tabla 5 Outliers IPRIX

Se obtienen los valores atípicos que se comprueba de nuevo que se acumulan en el año 2008.

6.6 Índice de Precios de Importación de Productos Industriales

6.6.1 Definición

El Índice de precios de importación (IPRIM) mide mensualmente la evolución de los precios de productos industriales que son vendidos en el mercado interior y proceden del mercado exterior.

6.6.2 Datos

En este caso, al igual en los Precios de Exportación los datos comienzan en Enero de 2005 hasta hoy, y tienen una toma mensual. La gráfica, que se puede ver que se comporta de forma similar a la de Precios de Exportación es la siguiente:

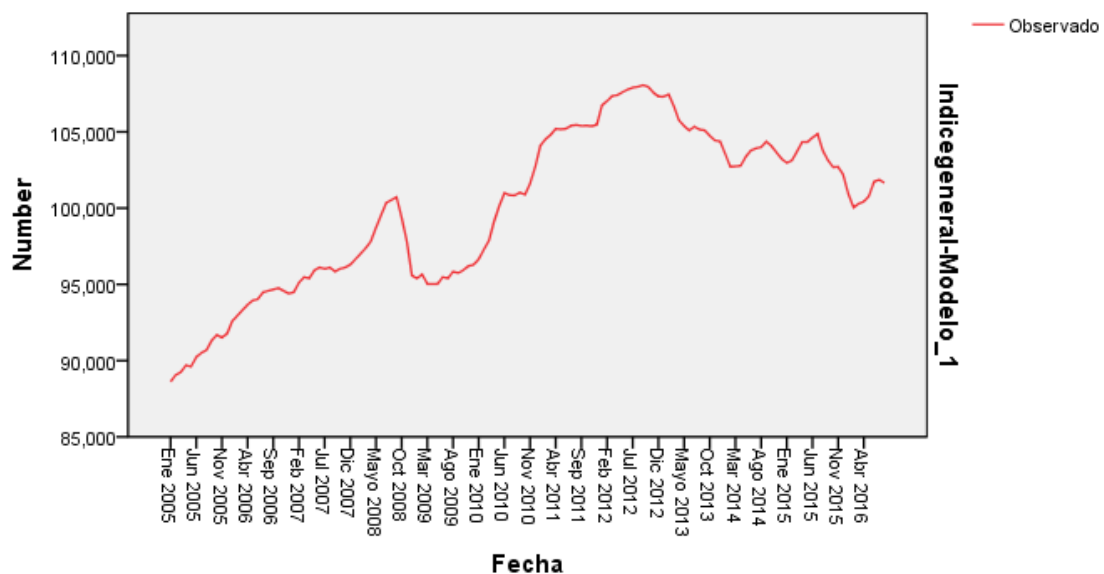


Figura 45 Gráfica del Índice de precios de importación de productos industriales (IPRIM)

Obtenemos la FAS y la FAP de los datos, que son las siguientes:

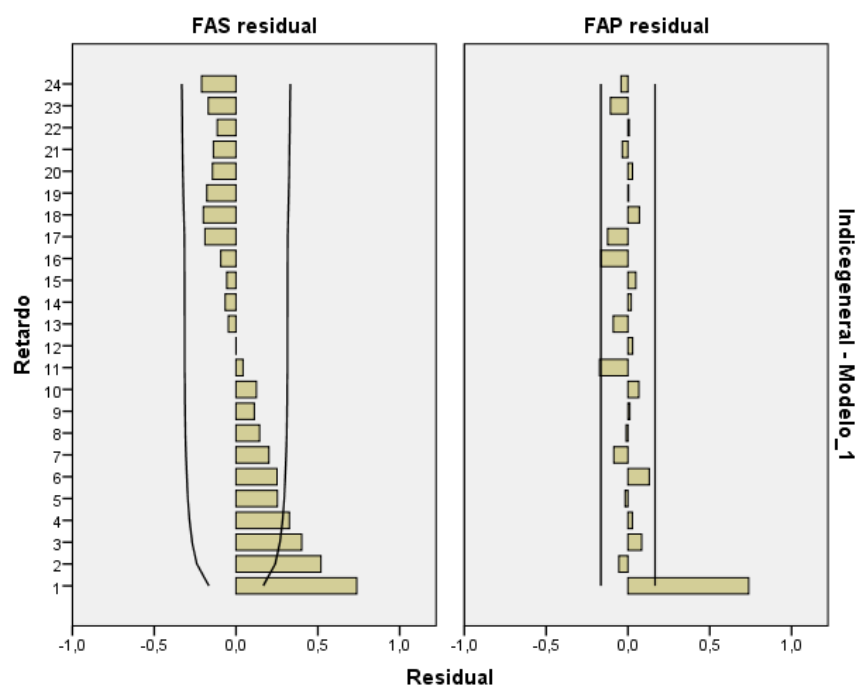


Figura 46 FAS y FAP del IPRIM

6.6.3 Análisis de datos

Analizaremos los datos planteando el modelo igual que el realizado anteriormente, es decir:

$$(1 - \phi_1 B) \cdot \nabla z_t = a_t \quad (70)$$

Obteniendo los siguientes resultados:

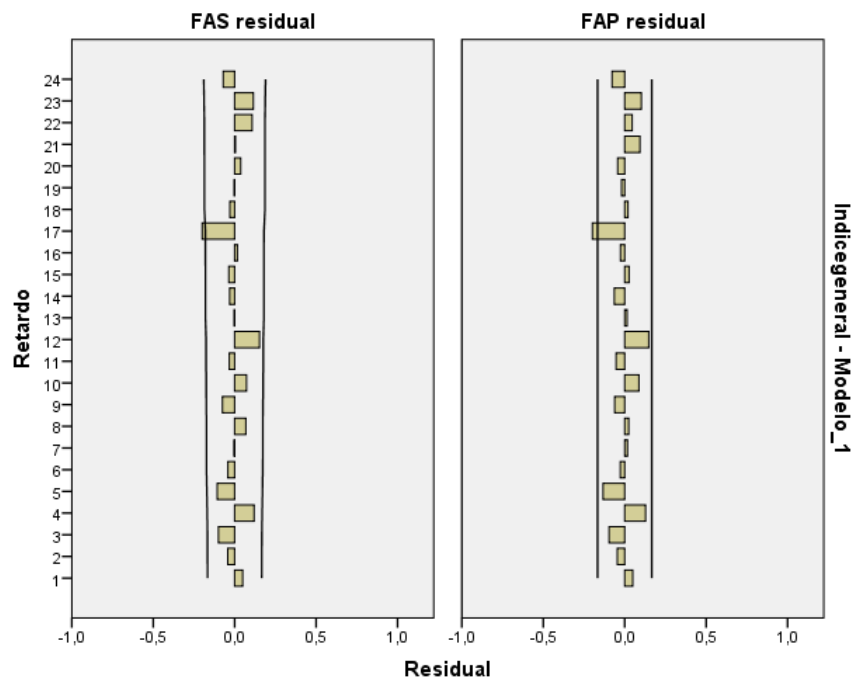


Figura 47 FAS y FAP del IPRIM tras aplicar el modelo ARIMA(1,1,0)

Se comprueba ahora si el modelo es correcto.

$$\begin{matrix} (1 - 0,546B) \cdot \nabla z_t = a_t + 0,126 \\ (7,652) \qquad \qquad \qquad (0,786) \end{matrix} \quad (71)$$

$$p - valor = 0.322 \quad (72)$$

Por todo lo mencionado en modelos anteriores podemos afirmar que el modelo ha obtenido correctamente la estructura de los datos. Por lo que no podemos rechazar la hipótesis nula de que los residuos son ruido blanco.

6.6.4 Outliers

No hay.

6.7 Tabla de Outliers

Periodos	Número outliers	IPI	Paro total	IPrI	IPRIX	IPRIM
Anterior a 2008	1		Febrero (2002) Aditivo			
2008	8	Febrero Tendencia lineal Junio Aditivo	Abril Tendencia lineal Octubre Tendencia lineal	Julio Aditivo Noviembre Innovativo	Octubre Innovativo Diciembre Cambio de nivel	
2009	3	Abril Tendencia lineal Agosto Innovativo				
2010	0					
2011	0					
2012	1				Enero Cambio de nivel	
2013	3			Marzo Innovativo Abril Aditivo Diciembre Aditivo		
2014	2			Febrero Cambio de nivel Agosto Aditivo		
2015	2			Agosto Aditivo	Agosto Innovativo	

2016	4			Enero Innovativo Junio Cambio de nivel Octubre Innovativo Diciembre Innovativo		
2017	1			Enero Aditivo		

Tabla 6 Outliers de todas las series temporales

Como se puede ver en la tabla superior los valores atípicos se concentran en el año 2008, por lo que se comprobará numéricamente la hipótesis de que los valores atípicos se presentan en mayor cantidad en los años en los que se dio un comportamiento extraño de las series temporales, para ello se comprobará que porcentaje de los mismos se encuentra en esos años.

Periodos	Número outliers	Porcentaje Outliers	Acumulado
Anterior a 2008	1	4	4
2008	8	32	36
2009	3	12	48
2010	0	0	48
2011	0	0	48
2012	1	4	52
2013	3	12	64
2014	2	8	72
2015	2	8	80
2016	4	16	96
2017	1	4	100

Tabla 7 Porcentaje de Outliers por periodo

Se observa que la hipótesis es correcta, puesto que en tan solo en el año 2008 se encuentran el 32%. Esto implica que un tercio de los datos se encuentran en un solo año de los más de diez analizados. Si se amplía al año 2009 se observa como en dos años se alcanza prácticamente el 50% de los valores atípicos obtenidos.

7. Estudio de los efectos de la crisis económica de 2008 sobre el paro con modelos ARIMA estacionales con Outliers

En este capítulo se va a analizar en profundidad la evolución del paro registrado en España y los efectos de la crisis económica. Para ello se tomará el modelo que se ha estimado en la sección 6.3 para el número de parados mensual en España. El procedimiento será el siguiente:

1. Se obtienen las previsiones a partir del último periodo observado que es Mayo de 2017. Por tanto obtendremos previsiones a partir de Junio de 2017.
2. Vamos a estudiar el efecto de la crisis económica mediante la capacidad predictiva del modelo. Para ello daremos un primer corte en Agosto de 2007. Es decir, tomaremos los datos desde Enero de 1996 hasta Agosto de 2007 y utilizando el modelo calcularemos previsiones a partir de ese momento. Ahí todavía no ha empezado la crisis económica, por lo que las previsiones nos darán lo que hubiera sido la evolución del paro si no hubiera habido crisis económica. Comparando estas previsiones con los valores reales obtenidos podremos visualizar el impacto de la crisis.
3. El mismo procedimiento se va a utilizar con punto de corte Septiembre del 2008 que es la quiebra de Lehman Brothers y en tercer lugar haremos lo mismo en Junio de 2011.

7.1 Previsiones de paro a partir de Mayo de 2017

Utilizando el modelo analizado en 6.3.3:

$$\nabla^2 \nabla_{12} z_t = \underbrace{(1 - 0,694B)}_{(15,930)} \cdot \underbrace{(1 - 0,372B^{12})}_{(0,219)} a_t + 68,308 \quad (73)$$

Se obtiene la Figura 48 en la cual se puede ver como las previsiones de paro continúan bajando. Esto indica que las previsiones son buenas, puesto que en el caso de que la serie siga comportándose como en los últimos periodos recuperará el empleo destruido por la crisis.

También ha de mencionarse que las previsiones tienen un intervalo de confianza que según avanza el tiempo tras el último dato real se va haciendo mayor, por lo que las previsiones empeorarán.

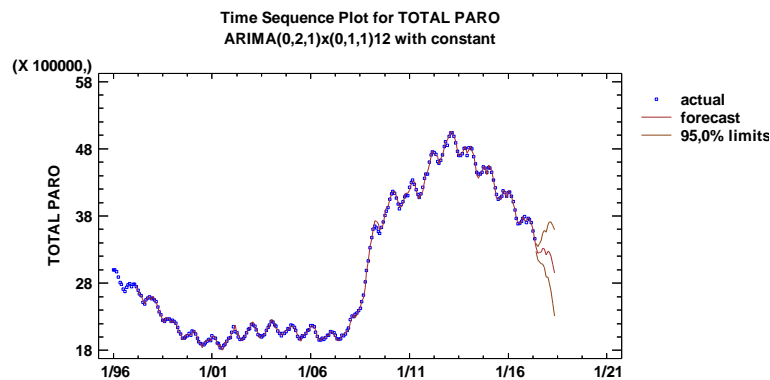


Figura 48 Previsiones Paro Mayo 2017

En la Tabla 8 se puede observar lo mencionado en la Figura 48, los límites superior e inferior se van alejando del valor previsto conforme avanzan las previsiones.

		<i>Límite 95,0%</i>	<i>Límite 95,0%</i>
<i>Fecha</i>	<i>Previsiones</i>	<i>Inferior</i>	<i>Superior</i>
6/17	3,33534E6	3,38323E6	3,28745E6
7/17	3,24909E6	3,15389E6	3,3443E6
8/17	3,25209E6	3,11553E6	3,38865E6
9/17	3,26507E6	3,08404E6	3,44609E6
10/17	3,31673E6	3,08803E6	3,54543E6
11/17	3,31062E6	3,03112E6	3,59013E6
12/17	3,223E6	2,88967E6	3,55633E6
1/18	3,27746E6	2,88741E6	3,66751E6
2/18	3,26209E6	2,81255E6	3,71162E6
3/18	3,20146E6	2,68978E6	3,71314E6
4/18	3,07757E6	2,50118E6	3,65396E6
5/18	2,95106E6	2,30749E6	3,59462E6

Tabla 8 Previsiones Paro Mayo 2017

Finalmente se comparará el resultado previsto para el mes de Junio de 2017 con el dato real.

$$\begin{aligned}
 \text{Predicción} &= 3.335.339 \\
 UCL &= 3.383.227 \\
 LCL &= 3.287.451
 \end{aligned}
 \tag{74}$$

7.2 Primer punto a analizar: Agosto 2007

El primer punto que se analizará será Agosto de 2007. Dicho punto se considera de gran interés en la crisis económica del 2008 porque es el punto previo a la misma. Sin embargo como se verá al terminar esa sección el modelo ya ha detectado el inicio de la crisis. A lo largo de este mes se hace patente el problema a nivel global de las hipotecas subprime, avisando el Banco Central Europeo de que se avecinaba una crisis financiera debida al capital riesgo y las hipotecas subprime inyectando los bancos liquidez en el sistema financiero. Hasta este momento no se observan grandes variaciones en la serie planteada, por lo que el modelo no puede prever el comportamiento extraño que tendrá posteriormente la serie. Hasta ese momento la serie de Paro total es la siguiente:

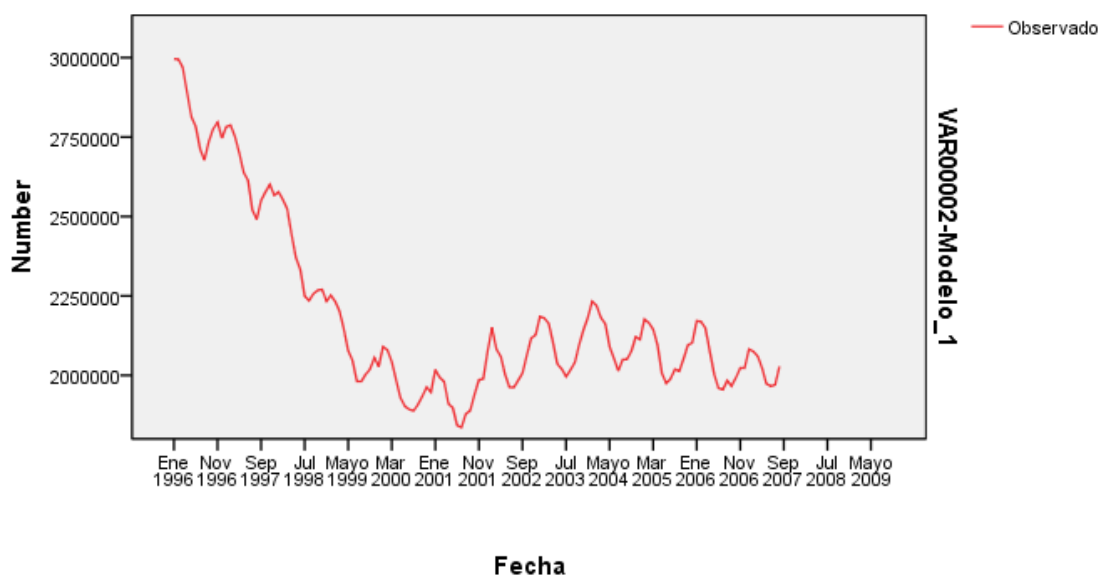


Figura 49 Gráfica Paro total (Agosto 2007)

Analizamos ahora dicha serie para obtener su modelo ARIMA con Outliers. Dicha serie se ha de analizar de nuevo a pesar de haberla analizado con anterioridad puesto que los datos incluidos en la serie completa podrían cambiar la bondad de ajuste del modelo. Utilizamos un modelo ARIMA(0,2,1)(0,1,1)₁₂ cuyo modelo y p-valor son:

$$\nabla^2 \nabla_{12} z_t = (1 - \theta_1 B) \cdot (1 - \Theta_1 B^{12}) a_t \quad (75)$$

$$p - \text{valor} = 0,762 \quad (76)$$

Se comprueba por lo tanto que el modelo planteado para la serie completa sigue siendo válido para la analizada hasta este instante, puesto que el valor del p-valor es mayor que 0,05.

$$\nabla^2 \nabla_{12} z_t = (1 - 0,84B) \cdot (1 - 0,423B^{12}) a_t + 69,203 \quad (77)$$

(15,184) (4,714) (0,410)

Los valores atípicos del modelo serían los siguientes:

			Estimación	SE	T	Sig.
TOTAL PARO-Modelo_1	Mayo 2001	Aditivo	38051,902	9911,635	3,839	,000
	Feb 2002	Aditivo	66451,016	9945,967	6,681	,000

Tabla 9 Outliers Paro total (Agosto 2007)

Los valores atípicos obtenidos se encuentran en los años 2001 y 2002. Dichos outliers pueden ser debidos a la breve recesión a comienzos de la década de los 2000 que se mencionaba en el análisis del sector industrial. Para el modelo las previsiones obtenidas con sus respectivos límites de predicción son las siguientes:

Predicción				
Modelo		Predicción	UCL	LCL
TOTAL PARO-Modelo_1	Sep 2007	2031178	2066785	1995570
	Oct 2007	2075726	2130251	2021200
	Nov 2007	2124760	2196732	2052787
	Dic 2007	2140300	2229482	2051118
	Ene 2008	2215528	2322103	2108953
	Feb 2008	2222224	2346557	2097891
	Mar 2008	2217222	2359769	2074675
	Abr 2008	2184468	2345731	2023205
	Mayo 2008	2137654	2318158	1957150
	Jun 2008	2130565	2330845	1930286
	Jul 2008	2145053	2365646	1924460
	Ago 2008	2204797	2446240	1963355
	Sep 2008	2216438	2488037	1944838
	Oct 2008	2269813	2572384	1967242
	Nov 2008	2327743	2662114	1993373
	Dic 2008	2352250	2719253	1985246
	Ene 2009	2436513	2836979	2036047
	Feb 2009	2452313	2887064	2017562
	Mar 2009	2456484	2926335	1986634
	Abr 2009	2432973	2938727	1927220
	Mayo 2009	2395471	2937920	1853022
	Jun 2009	2397764	2977688	1817840
	Jul 2009	2421701	3039869	1803533
	Ago 2009	2490966	3148135	1833786

Tabla 10 Predicciones Paro total (Agosto 2007)

La gráfica obtenida para los valores que se han previsto mediante el modelo es la siguiente:

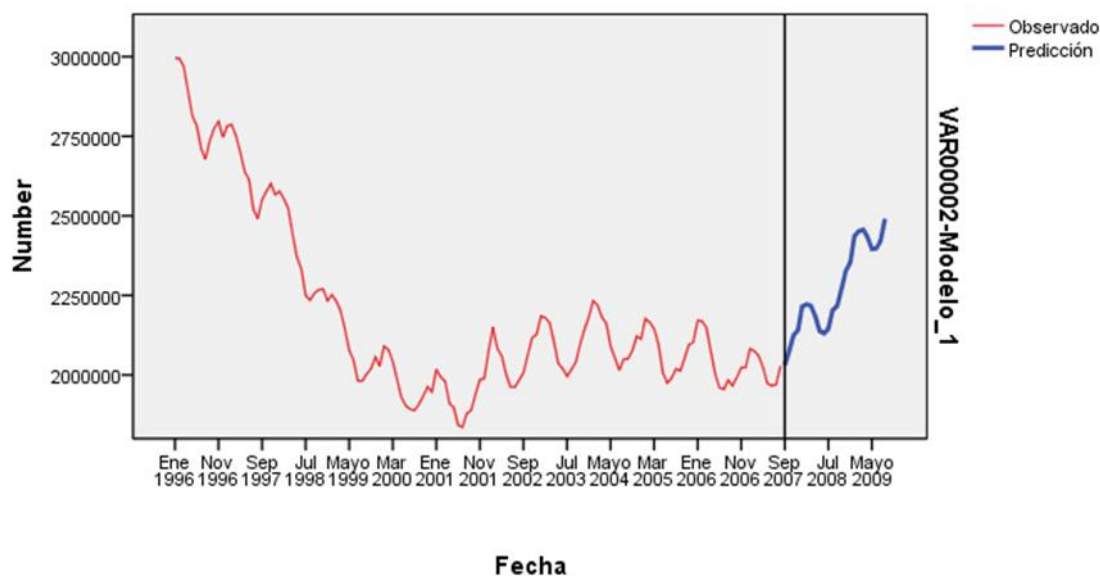


Figura 50 Previsiones paro total (Agosto 2007)

Como puede observarse, a pesar de que aparentemente no se ve el inicio de la crisis, el modelo ya lo ha detectado. Sin embargo el efecto real de la crisis fue mucho mayor. Para comprobarlo podemos comparar los resultados obtenidos con aquello que posteriormente ocurrió:

	Real	Previsión
Sep 2007	2017363	2031178
Oct 2007	2048577	2075726
Nov 2007	2094473	2124760
Dic 2007	2129547	2140300
Ene 2008	2261925	2215528
Feb 2008	2315331	2222224
Mar 2008	2300975	2217222
Abr 2008	2338517	2184468
Mayo 2008	2353575	2137654
Jun 2008	2390424	2130565
Jul 2008	2426916	2145053
Ago 2008	2530001	2204797
Sep 2008	2625368	2216438
Oct 2008	2818026	2269813
Nov 2008	2989269	2327743
Dic 2008	3128963	2352250
Ene 2009	3327801	2436513
Feb 2009	3481859	2452313
Mar 2009	3605402	2456484
Abr 2009	3644880	2432973
Mayo 2009	3620139	2395471
Jun 2009	3564889	2397764
Jul 2009	3544095	2421701
Ago 2009	3629080	2490966

Tabla 11 Datos reales frente a previsión Paro total (Agosto 2007)

La Figura 50 muestra la serie real hasta Agosto de 2007 y las previsiones que el modelo da a partir de ese momento. Como puede verse el modelo ha detectado el inicio de la crisis económica y las previsiones son de aumento de paro. Sin embargo no pudo prever la magnitud del incremento del paro. Se puede comprobar que a medida que pasa el tiempo los datos reales se alejan de la previsión a causa de la crisis. Por ello situaremos la serie real sobre la prevista con el fin de comprobar cómo la realidad se comportó distinto a lo que se podía haber previsto. Con el fin de que sea lo más claro posible se ha ampliado el intervalo temporal de la predicción. De esta forma ha sido posible ver como la serie real pierde a lo largo de 2008 y 2009 su componente estacional. Esto es indicativo de la gravedad de la crisis en un país como España, puesto que como se ha visto en el estudio inicial del sector industrial español el sector servicios hace que la estacionalidad afecte en gran medida al paro.

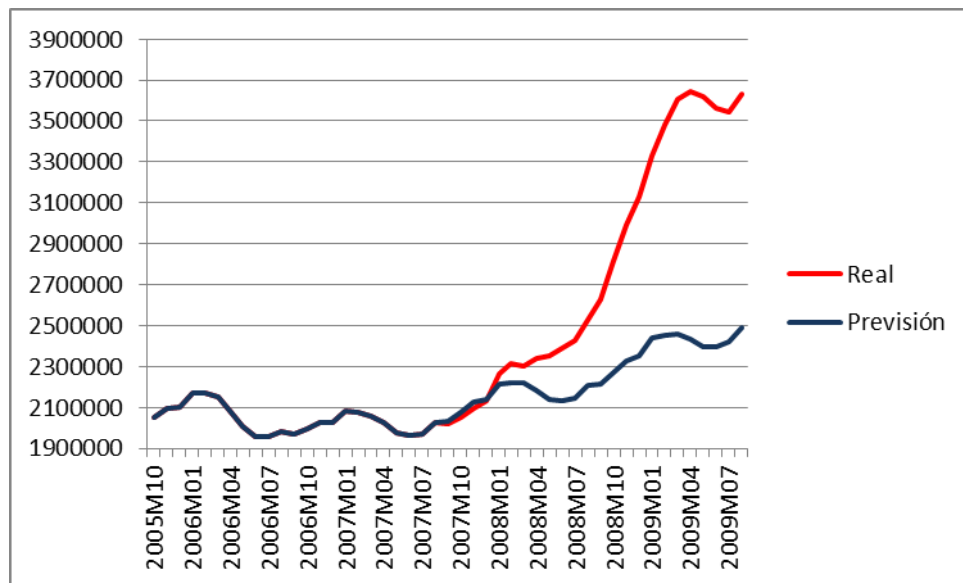


Figura 51 Evolución datos reales frente a previsión (Agosto 2007)

Para comprobar lo que verdaderamente se han alejado las previsiones de los datos reales calculamos las desviaciones porcentuales de la siguiente manera poniendo como ejemplo el primer dato, Septiembre 2007:

$$\left| \frac{e_t}{x_t} \right| \cdot 100 = \left| \frac{2017363 - 2031178}{2017363} \right| \cdot 100 = 0,68480487 \quad (78)$$

Calculamos el resto de valores de desviación porcentual:

Fecha	Desviación porcentual
Sep 2007	0,68480487
Oct 2007	1,32526139
Nov 2007	1,44604395
Dic 2007	0,50494307
Ene 2008	2,05121744
Feb 2008	4,02132568
Mar 2008	3,63989179
Abr 2008	6,58746547
Mayo 2008	9,17417121
Jun 2008	10,870833
Jul 2008	11,6140402
Ago 2008	12,853908
Sep 2008	15,5761021
Oct 2008	19,453795
Nov 2008	22,1300258
Dic 2008	24,823336
Ene 2009	26,7830919
Feb 2009	29,5688596
Mar 2009	31,8665713
Abr 2009	33,249572
Mayo 2009	33,8293088
Jun 2009	32,7394485
Jul 2009	31,6694107
Ago 2009	31,360951

Tabla 12 Desviación porcentual Previsto/Real Paro total (Agosto 2007)

La Tabla 12 muestra que en los primeros periodos el error es menor y conforme avanza el tiempo la desviación se hace notable.

7.3 Segundo punto a analizar: Septiembre 2008

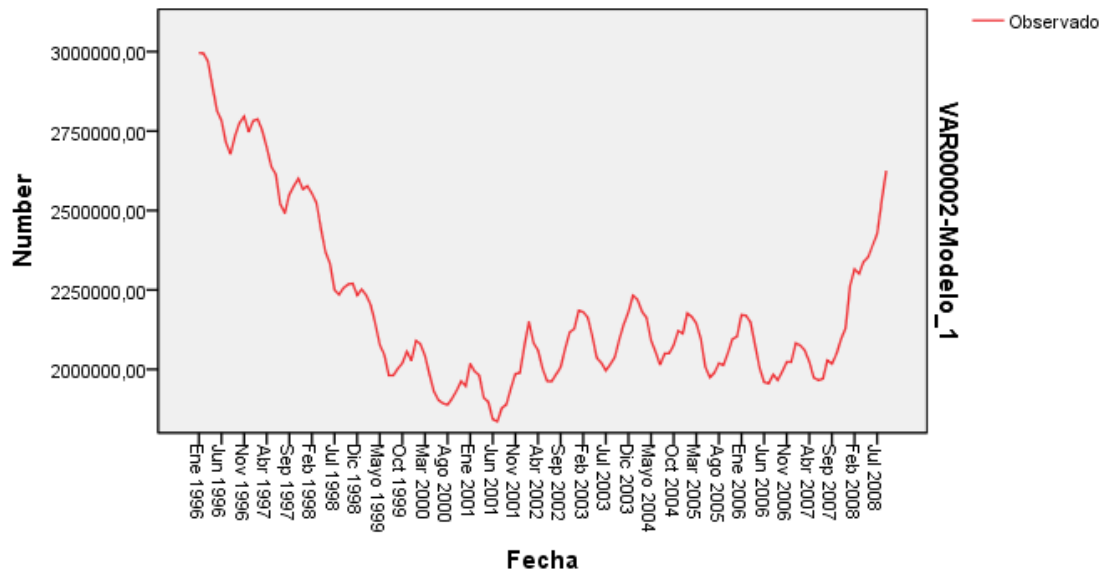


Figura 52 Gráfica Paro total (Septiembre 2008)

El segundo punto a analizar es septiembre de 2008. Esto se debe a que el día 15 de dicho mes se anunció la quiebra formal de Lehman Brothers, la cuarta mayor entidad bancaria de los Estados Unidos en aquel momento. Esta entidad llevaba varios semestres acumulando grandes pérdidas y perdiendo clientes. Dicha quiebra tuvo un gran impacto en los mercados del mundo entero, afectando también a España.

Al igual que en la primera serie analizada hemos de comprobar si el modelo utilizado sigue valiendo para la nueva serie. Por ello volvemos a analizar la serie del paro pero cortando los datos conocidos en septiembre de 2008. La estructura del modelo que se obtiene es la siguiente, que se puede comprobar que es la misma que en la serie del paro. Esto nos muestra de nuevo la robustez del modelo planteado para la serie del paro. Dicho modelo además arroja un p-valor mucho mayor que 0,05 por lo que se puede afirmar que ha atrapado bien la estructura de los datos:

$$p - \text{value} = 0,881 \quad (79)$$

$$\nabla^2 \nabla_{12} z_t = (1 - \theta_1 B) \cdot (1 - \Theta_1 B^{12}) a_t \quad (80)$$

$$\nabla^2 \nabla_{12} z_t = (1 - 0,858B) \cdot (1 - 0,446B^{12}) a_t + 65,834 \quad (81)$$

(17,660)(5,199) (0,460)

Observamos ahora los valores atípicos de la serie. Hay dos de los cuatro que revisten especial importancia, son los dos de 2008. Son importantes debido a que demuestran que el modelo está analizando correctamente los datos debido a que detecta una tendencia local en enero de 2008, que como hemos comprobado en el primer punto es ya un mes en el que se sufría los efectos de la crisis en el paro y finalmente un cambio de nivel en marzo.

Valores atípicos			Estimación	SE	t	Sig.
VAR00002-Modelo_1	Mayo 2001	Aditivo	38143,160	10156,901	3,755	,000
	Feb 2002	Aditivo	66298,009	10188,187	6,507	,000
	Ene 2008	Tendencia local	55991,486	10620,484	5,272	,000
	Mar 2008	Cambio de nivel	-64162,103	18066,807	-3,551	,001

Tabla 13 OutliersParo total (Septiembre 2008)

Con el modelo utilizado se obtienen previsiones de los siguientes periodos con el fin de compararlos con los datos reales.

	Modelo		
	VAR00002-Modelo_1		
	Predicción	UCL	LCL
Oct 2008	2726943	2762782	2691103
Nov 2008	2838705	2893097	2784313
Dic 2008	2929514	3000743	2858285
Ene 2009	3069852	3157495	2982209
Feb 2009	3135908	3239991	3031825
Mar 2009	3194755	3315506	3074004
Abr 2009	3234460	3372212	3096708
Mayo 2009	3255358	3410500	3100216
Jun 2009	3306107	3479060	3133154
Jul 2009	3366449	3557651	3175246
Agg 2009	3483776	3693674	3273878
Sep 2009	3572326	3801369	3343283
Oct 2009	3683063	3940070	3426055
Nov 2009	3804053	4089614	3518491
Dic 2009	3904155	4218898	3589413
Ene 2010	4053853	4398425	3709280
Feb 2010	4129334	4504396	3754272
Mar 2010	4197672	4603887	3791457
Abr 2010	4246934	4684966	3808902
Mayo 2010	4277454	4747965	3806944
Jun 2010	4337892	4841539	3834246
Jul 2010	4407988	4945422	3870555
Agg 2010	4535136	5107001	3963270
Sep 2010	4633572	5240506	4026638

Tabla 14 Predicción Paro total (Septiembre 2008)

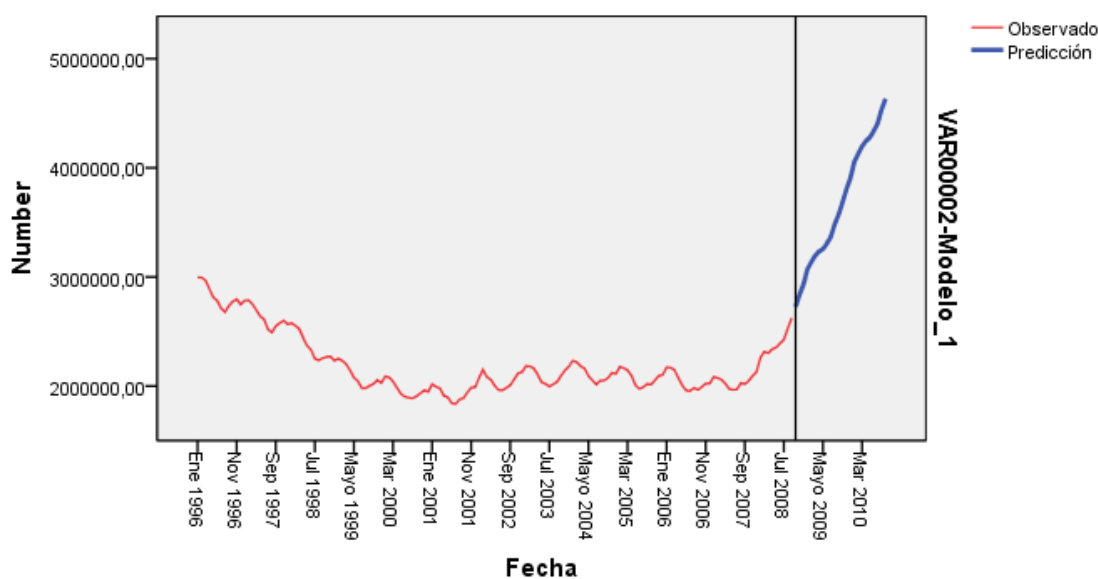


Figura 53 Previsiones paro total (Septiembre 2008)

Se comparan los datos reales con las previsiones y se obtiene la siguiente gráfica.

	Real	Previsión
Oct 2008	2.818.026	2726943
Nov 2008	2.989.269	2838705
Dic 2008	3.128.963	2929514
Ene 2009	3.327.801	3069852
Feb 2009	3.481.859	3135908
Mar 2009	3.605.402	3194755
Abr 2009	3.644.880	3234460
Mayo 2009	3.620.139	3255358
Jun 2009	3.564.889	3306107
Jul 2009	3.544.095	3366449
Ago 2009	3.629.080	3483776
Sep 2009	3.709.447	3572326
Oct 2009	3.808.353	3683063
Nov 2009	3.868.946	3804053
Dic 2009	3.923.603	3904155
Ene 2010	4.048.493	4053853
Feb 2010	4.130.625	4129334
Mar 2010	4.166.613	4197672
Abr 2010	4.142.425	4246934
Mayo 2010	4.066.202	4277454
Jun 2010	3.982.368	4337892
Jul 2010	3.908.578	4407988
Ago 2010	3.969.661	4535136
Sep 2010	4.017.763	4633572

Tabla 15 Datos reales frente a Previsión Paro total (Septiembre 2008)

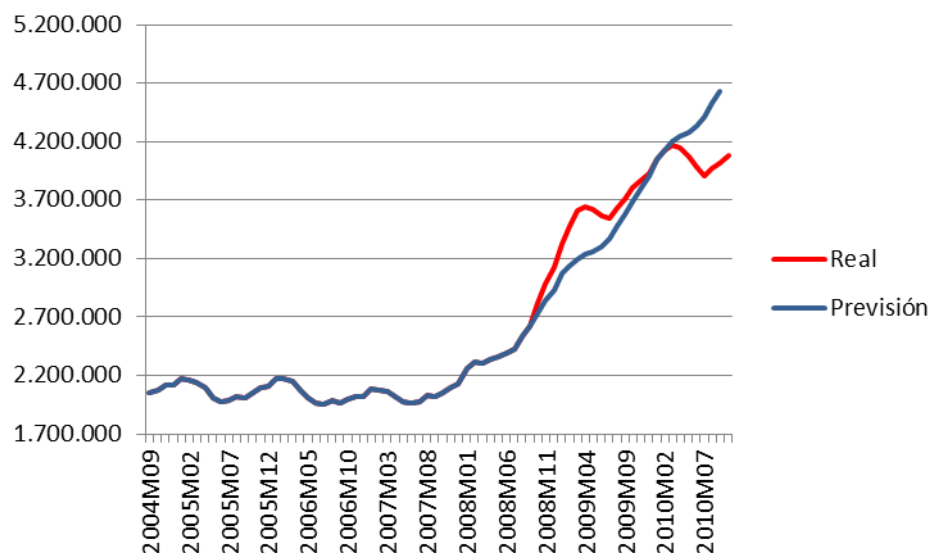


Figura 54 Evolución datos reales frente a previsión (Septiembre 2008)

Dicho gráfico aporta gran cantidad de información. Lo primero que se observa es que la pendiente de la recta de los datos reales es mayor que la de la prevista. Esto nos indica que a pesar de que los periodos anteriores a septiembre de 2008 habían sido malos en cuanto a

destrucción de empleo los siguientes iban a ser peores debido a, entre otros factores, la quiebra de Lehman Brothers. También es interesante comprobar cómo aunque la pendiente es mayor la recta real recupera la estacionalidad a principios de 2009, superando ya la etapa de crecimiento exponencial del paro, sin embargo la previsión había mantenido aunque muy levemente una cierta estacionalidad que no aumenta a principios de 2009. Calculamos las desviaciones porcentuales con el fin de comprobar la variación de la previsión frente a los datos reales de la siguiente manera poniendo como ejemplo el primer dato correspondiente a octubre de 2008:

$$\left| \frac{e_t}{x_t} \right| \cdot 100 = \left| \frac{2818026 - 2726943}{2818026} \right| \cdot 100 = 3,23215613 \quad (82)$$

Calculamos el resto de valores de desviación porcentual:

	Desviación porcentual
Oct 2008	3,23215613
Nov 2008	5,03681669
Dic 2008	6,37428439
Ene 2009	7,75133489
Feb 2009	9,93581302
Mar 2009	11,3897701
Abr 2009	11,2601787
Mayo 2009	10,0764363
Jun 2009	7,25918815
Jul 2009	5,01245029
Ago 2009	4,00387977
Sep 2009	3,69653482
Oct 2009	3,2898736
Nov 2009	1,67727851
Dic 2009	0,49566687
Ene 2010	0,13239494
Feb 2010	0,03125435
Mar 2010	0,7454256
Abr 2010	2,5228942
Mayo 2010	5,19531494
Jun 2010	8,92745221
Jul 2010	12,7772811
Ago 2010	14,2449192
Sep 2010	15,3271609

Tabla 16 Desviación porcentual Paro total (Septiembre 2008)

7.4 Tercer punto a analizar: Junio 2011

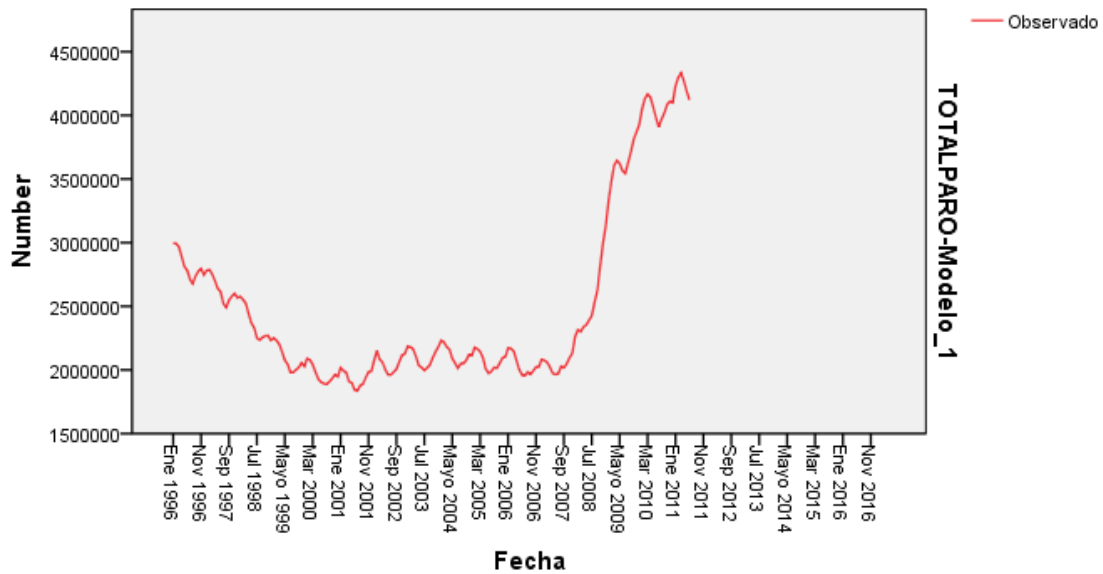


Figura 55 Gráfica Paro total (Junio 2011)

En tercer lugar se ha elegido septiembre del 2011 por ser una fecha en la que se aprecia claramente la estacionalidad y que, a priori, se podría suponer que la serie, tras haber sufrido un cambio de nivel permanecerá en el entorno actual. Analizamos la serie para obtener el modelo ARIMA que se comprueba que es el mismo que los utilizados en los puntos anteriores.

$$\nabla^2 \nabla_{12} z_t = (1 - \theta_1 B) \cdot (1 - \Theta_1 B^{12}) a_t \quad (83)$$

$$\nabla^2 \nabla_{12} z_t = \underset{(15,467)(5,384)}{(1 - 0,71B)} \cdot \underset{(0,232)}{(1 - 0,38B^{12})} a_t + 72,666 \quad (84)$$

$$p - valor = 0,523 \quad (85)$$

Obtenemos los valores atípicos en los cuales se puede observar que los puntos atípicos abundan en los años 2008 y 2009 como era de esperar. Es interesante observar como a partir de dicho año, es decir, desde junio de 2009 hasta el punto estudiado (junio de 2011) no hay ningún valor atípico. Esto se debe a lo comentado anteriormente; a partir de junio de 2009 la serie comienza a comportarse de forma normal.

Departamento de estadística
Estudio del Sector Industrial mediante modelos de función
de Transferencia (ARIMAX) con datos atípicos

Valores atípicos			Estimación	SE	t	Sig.
TOTAL PARO-Modelo_1	Mayo 2001	Aditivo	38398,004	10686,295	3,593	,000
	Feb 2002	Aditivo	66695,780	10705,253	6,230	,000
	Abr 2008	Tendencia local	61756,660	13474,151	4,583	,000
	Oct 2008	Tendencia local	57569,190	12575,536	4,578	,000
	Mayo 2009	Innovador	-84348,090	22035,449	-3,828	,000
	Jun 2009	Innovador	-119131,807	22160,083	-5,376	,000

Tabla 17 OutliersParo total (Junio 2011)

Obtenemos:

Predicción			
	Modelo		
	TOTAL PARO-Modelo_1		
	Predicción	UCL	LCL
Jul 2011	4046111	4087647	4004574
Agg 2011	4087547	4155331	4019763
Sep 2011	4116092	4210419	4021764
Oct 2011	4167460	4289622	4045298
Nov 2011	4181380	4332912	4029848
Dic 2011	4168719	4351208	3986231
Ene 2012	4270080	4485100	4055060
Feb 2012	4314506	4563596	4065416
Mar 2012	4321908	4606563	4037254
Abr 2012	4245614	4567282	3923946
Mayo 2012	4140834	4500919	3780749
Jun 2012	4040556	4440420	3640692
Jul 2012	3941419	4393839	3488998
Agg 2012	3959480	4467522	3451438
Sep 2012	3964723	4531218	3398228
Oct 2012	3992862	4620456	3365268
Nov 2012	3983626	4674812	3292439
Dic 2012	3947881	4705027	3190735
Ene 2013	4026230	4851595	3200865
Feb 2013	4047717	4943468	3151966
Mar 2013	4032254	5000478	3064030
Abr 2013	3933167	4975879	2890454
Mayo 2013	3805666	4924820	2686511
Jun 2013	3682740	4880232	2485248

Tabla 18 Predicción Paro total (Junio 2011)

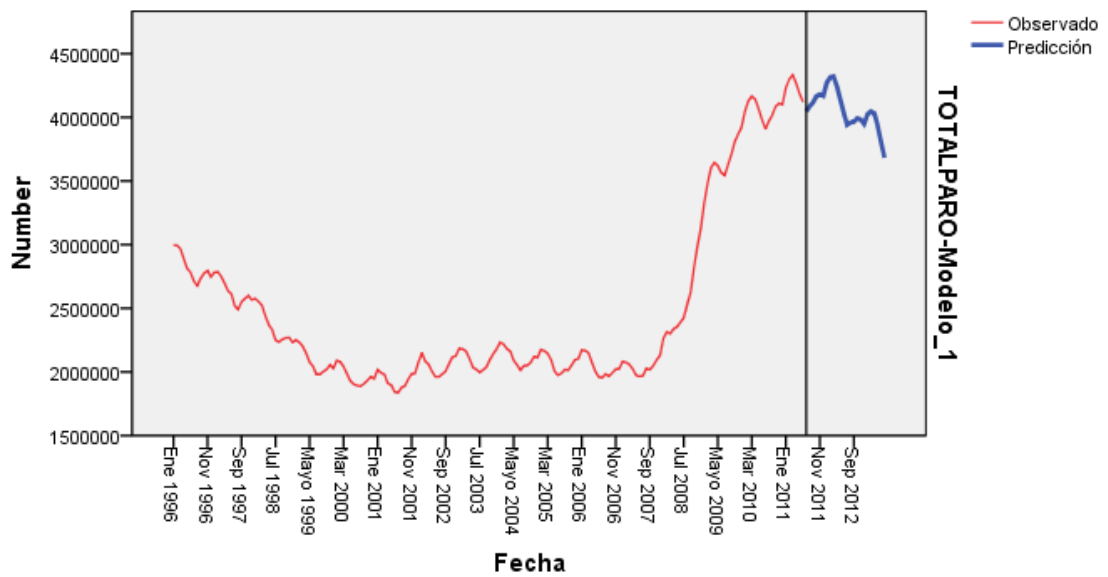


Figura 56 Previsión paro total (Junio 2011)

	Real	Previsión
Jul 2011	4079742	4046111
Agg 2011	4130927	4087547
Sep 2011	4226744	4116092
Oct 2011	4360926	4167460
Nov 2011	4420462	4181380
Dic 2011	4422359	4168719
Ene 2012	4599829	4270080
Feb 2012	4712098	4314506
Mar 2012	4750867	4321908
Abr 2012	4744235	4245614
Mayo 2012	4714122	4140834
Jun 2012	4615269	4040556
Jul 2012	4587455	3941419
Agg 2012	4625634	3959480
Sep 2012	4705279	3964723
Oct 2012	4833521	3992862
Nov 2012	4907817	3983626
Dic 2012	4848723	3947881
Ene 2013	4980778	4026230
Feb 2013	5040222	4047717
Mar 2013	5035243	4032254
Abr 2013	4989193	3933167
Mayo 2013	4890928	3805666
Jun 2013	4763680	3682740

Tabla 19 Datos reales frente a Previsión Paro total (Junio 2011)

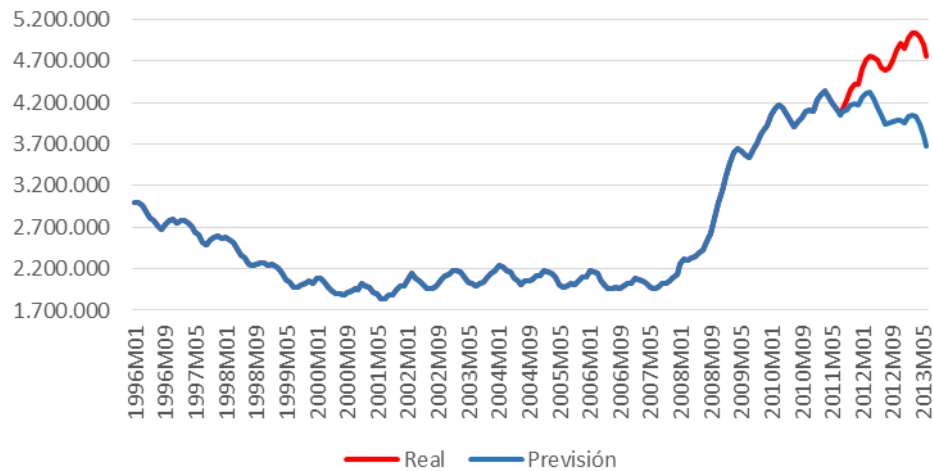


Figura 57 Evolución datos reales frente a previsión (Junio 2011)

Comparando los datos previstos y los reales se puede comprobar que la serie de datos sigue sufriendo impactos externos que impiden un comportamiento normal del modelo. Es por esto por lo que la previsión indica que el paro comenzará a descender levemente y respetando el efecto estacional, sin embargo los datos reales indican que continuó destruyéndose empleo. Habiendo obtenido la gráfica calculamos las desviaciones porcentuales para comprobar cuanto se han desviado las previsiones de los datos reales:

$$\left| \frac{e_t}{x_t} \right| \cdot 100 = \left| \frac{4079742 - 4046111}{4079742} \right| \cdot 100 = 0,82434134 \quad (86)$$

Calculamos el resto de valores de desviación porcentual:

	Con <u>Outliers</u>
Oct 2008	0,82434134
Nov 2008	1,05012749
Dic 2008	2,61790163
Ene 2009	4,43635136
Feb 2009	5,4085297
Mar 2009	5,7354005
Abr 2009	7,16872301
Mayo 2009	8,43768529
Jun 2009	9,02906775
Jul 2009	10,5100401
<u>Ago</u> 2009	12,1610769
<u>Sep</u> 2009	12,4524269
Oct 2009	14,0826668
Nov 2009	14,4013556
Dic 2009	15,7388329
Ene 2010	17,3922695
Feb 2010	18,8309996
Mar 2010	18,5789537
Abr 2010	19,1646365
Mayo 2010	19,6916922
Jun 2010	19,9193763
Jul 2010	21,1662688
<u>Ago</u> 2010	22,189286
<u>Sep</u> 2010	22,6912807

Tabla 20 Desviación porcentual Paro total (Junio 2011)

8. Conclusiones

En este proyecto se han analizado algunas series que aportan gran cantidad de información del sector industrial así como el número total de parados. Se han ajustado modelos ARIMA válidos y se han detectado los valores atípicos en todas estas series. El estudio de los outliers nos muestra la acumulación de los mismos en el año 2008. Antes de esa fecha no hay apenas outliers y con posterioridad hay muchísimos menos. La razón de la aparición de estos outliers en 2008 es evidentemente el impacto de la crisis económica. Este impacto era claramente apreciable en los gráficos de todas las series, pero los modelos con detección de outliers nos han permitido modelizarlos de forma conveniente. Las previsiones obtenidas son optimistas y en el corto plazo se prevé que continúe el crecimiento. Además, se ha hecho un estudio exhaustivo de la serie de paro registrado total en España. El impacto de la crisis sobre el paro ha sido enorme como es bien sabido. Como hemos obtenido un modelo válido, lo hemos utilizado para evaluar el impacto de la crisis en diferentes modelos sobre el paro. Así hemos podido observar que a pesar de haber previsto un aumento del paro en el año 2007, el impacto fue mucho mayor que el esperado. La previsión para los meses siguientes a Junio de 2017 es que continúe el descenso del paro.

Otra cuestión muy interesante es la robustez de los modelos ARIMA de la serie del paro. Se ha podido utilizar el mismo modelo ARIMA en todas las series generadas para los distintos puntos de interés analizados en el capítulo 7. También hemos visto cómo la detección automática de valores atípicos mejora las previsiones de los modelos. Las previsiones obtenidas por los modelos ARIMA en el corto plazo han mostrado ser muy precisas.

8.1 Posibles trabajos futuros

Se han planteado las siguientes líneas posibles de investigación:

- Inclusión de distintas series que tuvieron influencia en las distintas crisis tales como el precio mensual del barril de petróleo, Índice de Cifras de Negocio o Índice de Entrada de Pedidos con el fin de estudiar la relación entre ellas y la crisis del 2008.
- Estudio de las series de forma desagregada, diferenciando entre Bienes de consumo duradero, Bienes de consumo no duradero, Bienes de equipo y Bienes intermedios.
- Aplicación del método utilizado en el trabajo a distintos países de la Unión Europea con el fin de comparar los distintos sectores industriales y sus reacciones ante las distintas crisis.

9. Bibliografía

- [1]: Peña, D. (1993) *Modelos y métodos 2. Modelos lineales y series temporales*. Alianza.
- [2]: Peña, D. (2005) *Análisis de series temporales*. Alianza .
- [3]: Del Arce, R. (2009) *Modelos ARIMA* . Universidad Autónoma de Madrid, Departamento de Economía aplicada.
- [4]: Box, G.E.P. y Jenkins, G.M. (1977) *Time Series Analysis: Forecasting and Control*. Holden Day.
- [5]: Box, G.E.P. y Tiao, G. C. (1975): Intervention analysis with applications to economic and environmental problems", *Journal of American Statistical Association*, 70, 70-99.
- [6]: IBM (2012) *Manual del usuario del sistema del sistema básico de IBM SPSS Statistics 23*.
- [7]: Villagarcía, T. (2005) *Apuntes de series temporales*. Universidad Carlos III de Madrid, Departamento de Estadística, Leganés.
- [8]: Galeano, .P (2011) *Aspectos generales sobre la presencia de datos atípicos en series temporales*. Universidad Carlos III de Madrid, Departamento de Estadística, Leganés.
- [9]: Ministerio de Economía, Industria y Competitividad [en línea] Consultado el 25 de febrero de 2017, en <http://www.mineco.gob.es/>
- [10]: Instituto Nacional de Estadística (España). INEbase [en línea]. [Madrid]: INE. Consultado el 11 de febrero de 2017, en <http://www.ine.es/inebmenu/indice.htm>